Theory of Mind Development: Comparing Autism Spectrum Disorder Subgroups in Light of Changing Diagnostic Criteria

Michelle Robberts

RBBMIC002

A minor dissertation submitted in partial fulfilment of the requirement for the award of the degree of MA Research Psychology

Faculty of the Humanities

University of Cape Town

2010

COMPULSORY DECLARATION

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works, of other people has been attributed, and has been cited and referenced.

Signature:______________________________ Date:__________________________________

Acknowledgements

Thank you to all the children, parents and teachers who so generously gave of their time to assist with this project. To Susan Malcolm-Smith and Kevin Thomas, thank you for your guidance. And to Niel, Natalia, my parents and grandparents, thank you for your support and encouragement.

Table of Contents

List of Figures

List of Tables

Abstract

It has been proposed that autism is fundamentally a disorder of social relatedness. Severe deficits in theory of mind (ToM) - or the ability to understand that other people can have mental states different from our own and that these mental states influence behaviour – are commonly thought to explain the social-communicative deficits seen in autism spectrum disorders (ASDs). If deficits in ToM are responsible for the impairments found in ASD, these deficits should be found amongst all individuals with ASD (universality) and must be present throughout the course of the disorder (stability). It should also be possible to use ToM ability to devise a dimensional rating of 'level of functioning' within ASD. In light of ever greater prevalence estimates, it is imperative to understand the presentation of ToM deficits and their stability or development throughout the different subgroups of ASD in order to devise guidelines on nosology, intervention and management of the disorder. Additionally, the proposed changes to the Diagnostic and Statistical Manual of Mental Disorders (DSM), which will combine the different ASD diagnoses into one category, raise the need to reliably differentiate between ASD 'phenotypes'. Using a comprehensive ToM battery, I examined the universality and stability of ToM in order to establish whether ASD children of different ages and abilities could be differentiated into relatively distinct groups of ToM ability. This was done by looking at the presence of ToM skills in the current DSM diagnostic categories (low and high-functioning autism, Asperger's syndrome and pervasive developmental disorder not otherwise specified [PDD-NOS]), how these skills changed with age, and whether any obvious patterns of ToM ability emerged. The low and high-functioning autism and PDD-NOS groups displayed severely delayed ToM skills, with the low-functioning group scoring the lowest on ToM tests and the PDD-NOS group the highest. Surprisingly, the Asperger's syndrome group performed equivalently to the typically developing group on all the ToM tests. A cluster analysis showed three distinct clusters of ToM ability corresponding roughly with moderate to severe autism, high-functioning autism and Asperger's syndrome. Regarding stability of ToM over time, it was found that ToM develops at a rate comparable to typical development in high-functioning autism, PDD-NOS and Asperger's syndrome. The low-functioning autism group was the only group that did not display any obvious development in ToM. In terms of nosology, the results highlight major problems with the DSM-IV diagnostic criteria. Firstly, the PDD-NOS group did not have a uniform profile for ToM or general cognitive abilities. Secondly, while diagnostic problems plague the category of Asperger's syndrome, large differences in ToM ability were found within higher

functioning ASD, within which I include high-functioning autism, Asperger's syndrome and PDD-NOS. These results show the need to differentiate this very high functioning, relatively socially capable group from high-functioning cases with a more typical autism presentation. It is recommended that the diagnostic criteria be revised or that, if the ASD subgroups are to be grouped into a single autistic disorder category, dimensional categories within this diagnosis need to be made.

Introduction

The disorders that fall along the autism spectrum are characterized by impairments in communication and socialization, and the presence of restricted or repetitive behaviours (American Psychiatric Association, 2000). It has been proposed that autism is fundamentally a disorder of social relatedness (Kanner, 1943; Wing & Gould, 1979). To understand the extreme social aloofness, passivity in social interaction and oddness of interaction that is seen at varying levels throughout the autism spectrum, individuals with autism's cognitive understanding of social situations and their own and others' mental lives have been extensively studied (see, for example, Baron-Cohen & Swettenham, 1997). Severe deficits have been found in social cognition, or theory of mind (ToM), and these deficits have been proposed to underlie the social-communicative impairments seen in autism spectrum disorders (ASDs; Baron-Cohen, Leslie, & Frith, 1985).

What has received relatively little attention within the last three decades of theory of mind research is how the different subgroups within autism spectrum disorder compare with each other on ToM skills, and whether, or to what extent, ToM is able to develop within autism. Understanding the profile of ToM skills, and the development of these skills, in the different subgroups of ASD development is important from a theoretical perspective; the criteria for classification as a primary deficit are universality, specificity and stability (Ozonoff & McEvoy, 1994). Therefore, if deficits in ToM are central to autism, these deficits should be found amongst all individuals with autism, must be specific to autism, and must be present throughout the course of the disorder. ToM deficits are not specific to autism; these deficits are also found in individuals with schizophrenia and traumatic brain injuries, and in late-signing deaf individuals (Brüne & Brüne-Cohrs, 2006); however, it has been proposed that the *severity* of ToM deficits found in autism may be unique to the disorder. This study concerns the other two criteria for primacy: stability of ToM and universality within ASD. Once the universality of ToM and its development is known, it may be possible to use ToM ability to devise a dimensional rating of 'level of functioning' within ASD. Such a dimensional rating scale has immense practical significance: greater ToM skills have been associated with better social skills and decreased severity of autistic symptoms. ToM skills therefore have an impact on diagnosis, intervention and management of the disorder. Thus, the overarching goal of the study was to investigate whether children of different ages, abilities and ASD subgroups could be differentiated into relatively distinct groups of ToM ability. Such a level of functioning system will become particularly important if, as has been proposed, the various ASDs are grouped into a single autistic disorder category.

Review of the Literature

Autism Spectrum Disorder

ASD is an umbrella term that includes the pervasive developmental disabilities autism, Asperger's syndrome and pervasive developmental disorder not otherwise specified, as well as the relatively rare conditions Rett's disorder and childhood disintegrative disorder. These conditions are all characterized by (1) deficits in social interactions, (2) impaired verbal and non-verbal communication, and (3) repetitive, restricted or stereotyped behaviours, interests and activities (American Psychiatric Association, 2000). Notably, ASD is four times more likely to affect boys than girls (Kogan et al., 2009)

The text revision of the fourth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR; American Psychiatric Association, 2000) defines *autistic disorder*, or autism, as a developmental disorder whose onset must be prior to 3 years of age and in which all three core aspects of ASD must be present. Individuals with autism can be divided into high-functioning and low-functioning groups. Low-functioning autistic individuals have an IQ of below 70, and therefore fall in the intellectually disabled range, while high-functioning autistic individuals have an IQ higher than 70.

Individuals with *Asperger's syndrome* are similar to those with high-functioning autism in that they show impaired social interaction and communication, along with restricted and repetitive behaviour, and average intelligence. However, for a diagnosis of Asperger's syndrome to be made, there must be no delay in the onset of language abilities. *Rett's syndrome* is characterised by apparently normal development up until the age of 5 months, at which time there is a deceleration of head growth, loss of previously acquired purposeful hand skills and social engagement, and severe to profound psychomotor retardation with severely impaired language and poor coordination. *Childhood disintegrative disorder* is characterised by apparently normal development for at least the first 2 years, followed by a clinically significant loss of previously acquired skills before the age of 10 years in at least two of the following areas: language, social skills or adaptive behaviour, bowel or bladder control, play and motor skills. As part of the autism spectrum, this disorder features social and communicative impairments along with the presence of restricted, repetitive or stereotyped behaviour patterns, interest or activities. However, it cannot be better accounted for by another specific pervasive developmental area or Schizophrenia. Individuals with *pervasive developmental disorder-not otherwise specified (PDD-NOS)* show severe impairments in social interaction, associated with verbal or non-verbal communication deficits or repetitive and stereotyped behaviours, but do not meet the criteria for any specific pervasive developmental disorder. This may be because of the late onset of these deficits, or because of symptomatology that is atypical or subthreshold. For the full criteria for autism and Asperger's syndrome, see Appendix A.

The prevalence of autism spectrum disorders in South Africa is not known. A review done in 2003 estimated that the prevalence of ASD internationally is "at least 27.5 in 10 000" (Fombonne, 2003, p. 373). A recent study has reported a prevalence rate in the US of 110 in 10 000 (Kogan et al., 2009). The prevalence of autism is thought to be between 10 and 30 children in 10 000 (Baird et al., 2006; Fombonne, 2003; Kogan et al., 2009), with low-functioning autism making up approximately 70 percent of these cases and 30 percent of the entire ASD population (Chakrabarti & Fombonne, 2005). The rate of Asperger's syndrome is estimated at 2.5 in 10 000. The majority of the remaining children with ASD have PDD-NOS as Rett's syndrome and Childhood disintegrative disorder are extremely rare (both 0.2 – 0.4 in 10 000; Fombonne, 2003; Kozinetz et al., 1993). These alarmingly high rates highlight the need to understand the causes as well as the cognitive and behavioural aspects of ASD to inform intervention and management strategies. In particular it is important to include PDD-NOS as, although it is by far the most prevalent of the ASDs, it is – besides childhood disintegrative disorder – also the least studied (Matson & Boisjoli, 2007).

Changes to the DSM classification of ASD. In light of the upcoming release of the fifth edition of the DSM, the proposed changes to the diagnostic categories falling within the autism spectrum, and the reasons for those changes, will briefly be discussed. Recent studies have found considerable disagreement between clinicians' diagnoses of autism, Asperger's syndrome and PDD-NOS. For instance, in a survey of 466 professionals reporting on 348 cases, Williams et al. (2008) found that 44% of children given Asperger's syndrome, PDD-NOS, atypical autism, or "other ASD" labels actually fulfilled the DSM-IV criteria for autism. Differentiating between high-functioning autism and Asperger's syndrome has proven to be especially controversial. The main problems of the current diagnostic system will firstly be discussed. Secondly, the proposed changes to the ASD diagnostic categories in the DM-V will be outlined.

One of the problems of the DSM-IV diagnostic categories is the similarity between high-functioning autism and Asperger's syndrome. The differentiating factor between these two diagnoses, namely, the presence of a language delay, is difficult to establish when making a diagnosis in late childhood or even adulthood, and many have argued that this criterion is not a reliable indicator of aetiology, current presentation or outcome (Howlin, 2003; Prior et al., 1998; Swedo, Thorsen, & Pine, 2008). Furthermore, the argument has been raised that even in cases where language development is not delayed, it may not be normal – parents of no-speech-delay ASD children report more echolalia, pronoun reversal and pragmatic language deficits than parents of typically developing children (Eisenmajer et al., 1996). Even more problematic is that, although the earliest reports of Asperger's syndrome report lesser or different social impairments than those seen in autism, the degree of social-communicative impairment that is necessary for a DSM diagnosis of autism or Asperger's syndrome does not differ. This is highly problematic because while language should not be delayed in Asperger's syndrome, the DSM-IV does not exclude from a diagnosis of autism individuals with intact language development (as long as there are *communicative* impairments, which could be difficulties in initiating and sustaining conversation). As a diagnosis of autism takes precedence over a diagnosis of Asperger's syndrome, some researchers have argued that a diagnosis of Asperger's syndrome is therefore almost never valid as these individuals meet the social-communicative impairment requirements for a diagnosis of autism (Mayes, Calahoun, & Crites, 2000).

Rett's disorder, childhood disintegrative disorder and PDD-NOS are similarly problematic categories. Rett's disorder patients often have autistic symptoms for only a brief period during early childhood. Furthermore, the cause of Rett's disorder, a very specific genetic mutation, is now known (Amir et al., 1999) and therefore its inclusion in a manual for behavioural symptoms is no longer valid. The American Psychiatric Association has therefore concluded that the label of ASD is not appropriate for most individuals with Rett's disorder (American Psychiatric Association, n.d.). Childhood disintegrative disorder was included in the DSM-IV to stimulate research in this area. However, research has not succeeded in finding a distinction between the presentation or outcome of childhood disintegrative disorder and autism with developmental regression. Furthermore, it has proven difficult to establish whether development was typical before the onset of the disorder (American Psychiatric Assocation, n.d.). Regarding the diagnosis of PDD-NOS, Matson and Boisjoli (2007) very accurately state that PDD-NOS "is often diagnosed by what it is not (not autism), as opposed to what it is" (p. 75). Besides the assertion that individuals with PDD-NOS should not meet the criteria for any other pervasive developmental disorder, there are no specific criteria for PDD-NOS in the DSM-IV and, as such, presentation may vary widely from severe but atypical disability to very mild presentation. Perhaps for this reason, research has not been forthcoming for this category. The problems in establishing well-defined and distinct diagnostic categories that would stimulate research have lead to major revisions in the classification of these pervasive developmental disorders in the DSM.

In addition to the problems with diagnostic criteria that have been discussed, research has not been able to confirm qualitative differences between the subgroups, even though quantitative differences in functioning and symptomatology can be seen (Macintosh & Dissanayake, 2006; Manjiviona & Prior, 1995; Meyer & Minshew, 2002; Miller & Ozonoff, 2000; Witwer & Lecavalier, 2008). In other words, though individuals with Asperger's syndrome may have better language skills and individuals with PDD-NOS fewer symptoms than those with autism, the cognitive profiles are similar throughout the ASD subgroups and all subgroups show the same triad of impairments. Therefore, early drafts of the newly revised DSM-V propose that Asperger's syndrome, PDD-NOS and childhood disintegrative disorder be subsumed into the existing autistic disorder category and Rett's disorder be removed from the DSM entirely (Swedo et al., 2008).

This study will use the DSM-IV categories autism, Asperger's syndrome and PDD-NOS. Because childhood disintegrative disorder and Rett's syndrome occur so infrequently in the population, and because Rett's syndrome is no longer considered to be part of ASD, these categories will not be studied. The goal in using the DSM-IV categories in this study (though they will soon not be used anymore) is to see whether the groups can be differentiated on ToM performance. If clear profiles for the different subgroups do not emerge on ToM testing after controlling for factors such as general intellectual ability, as have been found on other cognitive measures (for example, Manjiviona & Prior, 1995; Noterdaeme, Wriedt, & Höhne, 2009; Thede & Coolidge, 2007), this would lend support to the reclassification of ASD into a single category.

Furthermore it is foreseeable that the proposed diagnostic changes will impact on the educational and healthcare systems. Though having one diagnostic category will solve much of the confusion surrounding appropriate diagnoses and will emphasize the similar nature (and possibly aetiology) of the ASDs, it will create a new set of problems that need to be addressed. ASD as a whole is an extremely heterogeneous group, and a single category is too broad to be of use as an indicator of disability, outcome, or best treatment. As will be discussed, individuals with ASD vary widely in their intellectual ability, in their social interaction and communication skills, and in their abilities to perform activities of daily living. Some classification method will therefore be necessary to distinguish between individuals at different levels of severity of the disorder. I suggest that the next important goal in ASD research will be to identify homogenous subgroups not only for research purposes, but to permit meaningful educational placement and intervention strategies. This classification system should follow a dimensional rather than categorical approach as was previously used (autism, Asperger's syndrome, PDD-NOS, etc.). The benefits of a dimensional classification system are that it (1) acknowledges a continuum of skills, (2) allows growth, and (3) provides an indication of outcome.

A dimensional rating system for DSM-V autistic disorder has been proposed, as shown in Table 1, but the greatest problem with this proposed classification system is that a fixed way of measuring this has, to my knowledge, not yet been proposed. People's opinion of what is considered "substantial support" may differ between places, cultures or individuals. A fixed way of measuring level of functioning is needed to ensure that all classification is consistent across studies, clinics or continents.

Table 1.

Proposed Dimensional Ratings for DSM-V Autistic Disorder

Dimensional Ratings for DSM-V ASD	Social Communication	Fixated Interest and Repetitive Behaviours
Requires very substantial support	Minimal social communication	Affects most activities; resistance to interference
Requires substantial support	Some social communication but noticeably disrupted	Frequent, interfering
Less severe	Even with support, noticeable impairments	Occasional, some interference
Subclinical symptoms	Clear impairments but more positive skills	Odd or excessive but no interference
Normal variation	Maybe awkward or isolated but within normal limits	Within normal limits for developmental level and no interference

Note: Proposed DSM-V ratings as of January 2011 (C. Lord, personal communication, January 19th, 2011).

A different type of classification system using a measurable factor that can be standardised, for instance ToM, may be a better approach to the problem of establishing fixed, universally acknowledged dimensional categories. For ToM to be useful as an indicator of level of functioning it must be able to discriminate different groups within ASD, it must serve as an indicator of adaptive functioning, and it must have recognised norms for different developmental levels. ToM, its association with social and adaptive functioning, and its development in typically developing children and children with ASD will be discussed next.

Theory of Mind

ToM refers to the ability to understand that other people can "want, feel and believe things" (Baron-Cohen, Leslie, & Frith, 1985, p.38) and to recognize that these mental states influence people's behaviour. The false belief paradigm has predominantly been used to test ToM (Fodor, 1992), as it is a reliable and easy way to test whether participants understand the distinction between thoughts and the real world. The classic false belief test used by Wimmer and Perner (1983) and Baron-Cohen and colleagues (1985), the "Sally-Anne test", shows Sally placing a marble in a basket and leaving the room. While she is away, Anne removes the marble from the basket and hides it in a box. Participants are then asked, "When Sally returns, where will she look for the marble?" The question is answered correctly (i.e., "in the basket") if the participant understands that Sally's belief does not represent the reality of the situation. This understanding of other people's beliefs is called first-order belief attribution.

It is assumed that ToM should be positively correlated with social competence. After all, for social reciprocity it is necessary to understand that other people have mental lives which, even though they are not visible, constantly affect their actions, and that these mental lives may be different from our own (Dennett, 1989).

ToM and social competence. There is some evidence that individuals' ToM skills are related to their social competence. A series of studies have found associations between ToM and social competence in typically developing children (see Bosacki & Wilde Astington, 2001; Repacholi & Slaughter, 2003), and several studies with ASD participants found that increased ToM was related to better social skills and fewer ASD symptoms and maladaptive behaviours (Fombonne, Siddons, Achard, Frith, & Happé, 1994; Frith, Happé, & Siddons, 1994; Lerner, Hutchins, & Prelock, 2010). Similarly, Travis, Sigman and Ruskin (2001) found that joint attention and increased empathy were associated with lesser degree of impairment in social interaction in individuals with ASD. However, it seems that the relationship between ToM and social competence is a complex one: Frith et al. (1994) found that children with ASD who passed false belief tasks showed greater social insight, but no more simple sociability, and Joseph and Tager-Flusberg (2004) found that although ToM was correlated with communicative competence, it could not explain the variance in the severity of social impairments or repetitive behaviours in ASD. Hence, it is unclear whether ToM represents a higher-order elaboration of the fundamental empathic abilities that are impaired in autism, or whether it is an independent cognitive module that is necessary, but not sufficient, for social competence (see Leslie & Frith, 1990).

To summarise, the association between social skills or social reciprocity and ToM is not clearly delineated – it would seem that ToM is necessary for social skills, but does not guarantee social competence; perhaps because a variety of ToM skills are needed for social competence. Research has recently begun to address this by examining multiple aspects of ToM, such as understanding of desires, emotions, and intentions, as well as precursors to ToM, such as pretend play and joint attention. A developmental examination of ToM skills and its cognitive influences follows.

The development of ToM. As emotions, desires, beliefs and intentions are all internal states it is conceivable that these might all be equally hard to understand, or that children from different backgrounds, and with different sets of experiences, might understand some types of mental states before other children do. Yet studies of ToM development from around the world show that, for all people, some mental states are harder to understand than others, and point to a similar trajectory in the development of ToM (Wellman, Cross, & Watson, 2001; Wellman & Liu, 2004). This universal developmental trajectory is present even in traditional hunter-gatherer societies who have had very little contact with Western culture (Avis & Paul L. Harris, 1991).

ToM and its precursors develop from a very young age. From 14 to 24 months of age, typically developing children start to engage in pretend play and show joint attention (i.e., looking at an object a parent is looking at; U. Frith & C. D. Frith, 2003). Pretend play and joint attention are argued to be precursors to ToM (Leslie, 1987). The ability to engage in pretend play shows that the child is aware of mental concepts and the distinction between what is real and what is imaginary. Joint attention is critical for social reciprocity as it requires an understanding of another person's intent, and a desire for sharing interest or enjoyment (Charman et al., 2000). At around age 2 years, children spontaneously begin to talk about their own mental states, express desires, and show an understanding of other people's desires (Wellman & Woolley, 1990). Between 3 and 5 years old, children begin to understand others' false beliefs, differences between appearance and reality, and their own previous false beliefs (Bibby & McDonald, 2005; Naito, Komatsu, & Fuke, 1994). Between 5 and 7 years old, children are able to understand second-order beliefs: In other words, they grasp that a person can have beliefs about other people's beliefs (Perner & Wimmer, 1985; Sullivan, Zaitchik, & Tager-Flusberg, 1994). The difference between jokes and lies begins to be appreciated by 6-10-year-olds, as are language forms such as metaphor, sarcasm and irony (Ackerman, 1981; Brüne & Brüne-Cohrs, 2006; M. R. Pollio & H. R. Pollio, 1979). This is an important indicator of ToM ability, as these language forms require an understanding of the speaker's intent. From 9 to 11 years old, children are able to recognize social faux pas, for example saying to a parent, "You have a lovely daughter", when their child is in fact a boy (Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999). Although ToM is clearly a complex set of skills, which is likely to continue to develop throughout the lifespan, ToM development after this point has not been studied.

Cognitive factors associated with ToM development. The aim of this study is not to investigate the potential causes and correlates of ToM. However, it is important to note which factors are associated with ToM task performance, as these factors could confound results when working with different clinical groups who may have various impairments beside ToM.

Numerous studies have reported correlations between language development and ToM in typical and clinical samples (see Milligan, Astington, & Dack, 2007 for a review). Milligan et al. (2007) reported significant correlations between false belief performance and receptive vocabulary, semantic and syntactic knowledge, general language skills and memory for complements. It may be that these various linguistic skills are necessary for ToM to develop (Astington, 2001; Astington & Jenkins, 1999); conversely, that ToM is necessary for language learning (Sabbagh & Baldwin, 2001); or it may be that there is a bidirectional relationship between ToM and language development (Slade & Ruffman, 2005). Whatever the relationship between ToM and language may be, it is important to be aware of language skills when assessing ToM.

Various aspects of executive functioning have also been correlated with ToM performance. Executive functioning is an umbrella term that includes a wide array of skills such as planning, judgement, inhibitory control, mental flexibility and generativity (Schneider, Schumann-Hengsteler, & Sodian, 2005). Correlations between the executive functions inhibition, cognitive flexibility (set shifting), working memory, planning, verbal fluency and ToM performance have been found in typical development (Carlson & Moses, 2001; Carlson, Moses, & Hix, 1998; Hughes, 1998; Perner & Lang, 1999; Sabbagh, Xu, Carlson, Moses, & Lee, 2006) as well as in various clinical samples (Greig, G. J. Bryson, & Bell, 2004; Henry, Phillips, Crawford, Ietswaart, & Summers, 2006; Joseph & Tager-

Flusberg, 2004; Pellicano, 2007; Perner, Kain, & Barchfeld, 2002; Saltzman, Strauss, Hunter, & Archibald, 2000; Wong, 2004). Proposed explanations for the observed correlation between executive functions and ToM include that both of these are linked to the development of other skills (e.g., language); that skills such as inhibition and working memory are necessary to reason about false beliefs as, for example, the belief needs to be held in mind and the desire to mention the real world state (rather than the character's false belief) needs to be suppressed; and that ToM is an important influence on the development of executive functions (Hughes, 1998). As many papers have noted deficits in executive function in autism, including deficits in shifting between instructions, generativity and processing speed (Ozonoff & Jensen, 1999; Ozonoff & McEvoy, 1994; Perner et al., 2002; Verté, Geurts, Roeyers, Oosterlaan, & Sergeant, 2006; Wong, 2004), these skills are again important to be aware of when assessing ToM.

ToM still remains a problematic concept – much like ASD it is an umbrella term which includes a variety of mental concepts, which have been poorly defined in the literature. This applies especially to the 'upper' and 'lower' ends of ToM – what would be the easiest task that still requires imputation of mental states, and on the other hand, when does ToM development end? At the same time, there are behaviours that are clearly essential for social reciprocity that do not necessarily require a ToM, such as making eye contact. Any attempts at examining social competence by measuring ToM must keep such conceptual problems in mind. It is also important to be aware of the fact that ToM, or at least ToM task performance, may be influenced by cognitive skills such as inhibition and verbal ability. It is unlikely that ToM has a general cognitive substrate. For example, some non-human primates are capable of tactical deception and display causal knowledge of predator behaviour (evidence for ToM), yet do not display higher or cognitive and verbal abilities (Byrne & Whiten, 1988; Whiten, 1997; Zuberbühler, 2000). However, ToM tests such as the false belief task have traditionally required that the participant be able to comprehend questions and respond appropriately, and hence completion of the task itself may have required verbal skills. By including a greater variety of ToM tests it is hoped that a more accurate picture of an individual's ToM may be obtained and that performance on a variety of ToM tests will be a good indication of social competence. Having explored ToM and its correlates in typically developing children, I now turn to the case of ToM in ASD.

ToM in ASD

It is thought that both the social aloofness and oddness in social interaction seen in ASD stem from an inability to empathise with others and to understand that others have mental states (Baron-Cohen et al., 1985; Wing, 1981). This theory has mostly been tested using the cognitive paradigm of ToM, and indeed, a plethora of research has found that individuals with ASD show severe deficits in developing ToM (Baron-Cohen et al., 1985; Holroyd & Baron-Cohen, 1993; Ozonoff & McEvoy, 1994; Tager-Flusberg, 2007). For instance, many children with ASD do not appear to understand the distinction between appearance and reality and are not good at recognizing mental state words such as “think” and “know”. They also tend not to engage in pretend play or imitation, and may not follow the gaze of a speaker (Charman et al., 2000). Although they may understand behaviour based on desires, children with ASD may struggle to understand complex causes of behaviour such as beliefs (Baron-Cohen & Swettenham, 1997). Furthermore, they tend to struggle with non-literal speech such as metaphor and irony (Happé, 1993). ASD children’s ToM difficulties cannot be attributed to low IQ, as children with Down’s syndrome have similar or lower IQ scores, but perform significantly better on false belief tests (Baron-Cohen et al., 1985). Thus, most ASD children show severe deficits in ToM ability that may be able to account for the social, communicative and imaginative deficits seen in ASD (Baron-Cohen & Swettenham, 1997). In order to account for these deficits, ToM needs to be both stable (i.e., it should never develop to normal levels while a diagnosis of ASD still applies) and universal throughout ASD. The stability of ToM in ASD throughout the lifespan will firstly be discussed.

Delayed or deviant development of theory of mind? Although most ASD individuals fail false belief tasks such as the Sally-Anne test, a significant proportion of autistic individuals (around 15 to 55 percent) *do* pass first-order false belief tests (Happé & Frith, 1996). These individuals are usually older and have a higher verbal mental age than autistic individuals who fail first-order false belief tasks (Happé & Frith, 1996; Ozonoff & McEvoy, 1994). This observation has led researchers to believe that individuals with ASD may have a specific developmental delay in ToM (Baron-Cohen, 1989; Baron-Cohen & Swettenham, 1997). This hypothesis predicts that ToM skills develop in ASD, and develop in the same order as in typical development, but are present only at a much later chronological and even mental age than is found in typical development and in intellectual disability without ASD (Baron-Cohen, 1989; Holroyd & Baron-Cohen, 1993). According to the theory of specific developmental delay, most individuals with ASD have ToM abilities equivalent to a 1-2 year-old typically developing child; that is, they do not show joint attention and imaginative play. Those individuals who display false belief reasoning skills are delayed at the next developmental level and therefore still display marked social and communicative impairments (Happé, 1994; Kaland et al., 2002), which are reflected in their difficulty in more naturalistic tests of social situations, such as the Faux Pas test (Baron-Cohen et al., 1999).

Supporting the delayed development hypothesis, Happé (1995) found that all ASD individuals with a verbal mental age of 11 years 9 months passed first-order false belief tasks. Given that ASD individuals who pass ToM tests are older and have a relatively high verbal ability, it follows that children with higher functioning ASD may show ToM development (albeit delayed), while low-functioning children never achieve false belief reasoning skills.

The delayed development hypothesis is also supported by the findings of a longitudinal study by Steele, Joseph, and Tager-Flusberg (2003), in which autistic children, aged 4 to 14 years, showed significant improvement in ToM abilities over the course of 1 year. Further evidence for developmental delay comes from Paynter and Peterson (2009), who found increased performance on ToM tasks with increasing age and mental ability (both verbal and nonverbal) in children with high-functioning autism and Asperger's syndrome.

In contrast to these results, longitudinal studies by Ozonoff and McEvoy (1994) and Holroyd and Baron-Cohen (1993) found no improvement in false belief scores of individuals with ASD at, respectively, 3 and 7 years after baseline testing. Their participants did not show increases in nonverbal mental ability either, although overall the group did show a significant increase in verbal ability (increased verbal skills were not seen in all participants, however). These results suggest a ceiling effect in the development of ToM, and therefore a deviance in development. One explanation for these contrasting results might be that Steele et al. (2003) used a more developmentally sensitive test battery, rather than only false belief tests: for instance, they included tests for early developing aspects of ToM, such as desire-based action tasks. Indeed, those researchers found that most of the improvement took place between early ToM and first-order ToM abilities, which could not have been measured by the false belief tests used in the other two studies. The results from Holroyd and Baron-Cohen (1993) and Happé (1995) also suggest an association between general cognitive ability and ToM task performance. Hence, it could be that Steele et al. (2003) and Paynter and Peterson (2009) used higher-functioning children in their samples, and that this was the reason for the improvement seen by these researchers.

Studies have also examined the *delay vs. deviance* argument by looking at developmental sequences within ToM performance. The delayed development hypothesis requires that a similar developmental trajectory is seen in ASD as in typical development; with pretend play preceding emotion attribution, which precedes false belief reasoning, and so forth. Sparrevohn and Howie (1995) found such a developmental progression in children with autism on tasks ranging from inferred belief to second-order false belief. Unfortunately, they did not have a control group to compare this performance to, and neither did they follow these children across time to see if their performance improved.

Two studies that examined performance on a range of ToM tasks support deviant rather than delayed development in autism. Peterson, Wellman, and Liu (2005) found that individuals with high-functioning autism, aged 6 to 14 years, showed a different ToM developmental pathway to both typically developing children and deaf children with a ToM deficit. Although all children showed the same developmental sequence for acquiring early ToM abilities, children with ASD found the false belief task more difficult than a hidden emotions task, while typically developing, late-signing deaf and native signing deaf children found the hidden emotions task the most difficult. So, although both autistic and deaf children have deficits in ToM, this study suggests that deaf children follow the same developmental pattern as typically developing children, whereas autistic children may follow a different developmental pattern, which may be unique to autism. These results on the hidden emotion task have not been replicated as yet.

Serra, Loth, van Geert, Hurkens, and Minderaa (2002) studied ToM development in 4-6-year-old children with PDD-NOS compared with controls matched on verbal and non-verbal mental age. They found that development of ToM abilities had taken place in the PDD-NOS group 6 months after baseline testing, but that this development was markedly slower than in the typically developing control group, and was not statistically significant. Furthermore, the developmental pattern between the groups differed considerably. While the typically developing group showed a relatively stable phase followed by rapid increase in ToM scores, the PDD-NOS group showed a pattern of increase, decrease, and increase again before stabilizing. Burack and Volkmar (1992) have noted that children with ASD are more likely to show developmental regressions than are typically developing children, and low-functioning autistic children are more likely to show developmental regressions than are high-functioning autistic children.

The studies on ToM progression in ASD thus far support both a delayed and deviant hypothesis for ToM development. ToM development seems to occur in at least some individuals with ASD, but it is slower and possibly follows a different developmental pattern to ToM in typically developing children. The results from the longitudinal studies by Holroyd and Baron-Cohen (1994) and Ozonoff and McEvoy (1994) further point to a possible plateau in ToM development in ASD.

In summary, at least some findings suggest that ToM deficits may not be as stable as was previously thought, and that ToM may develop in certain individuals or subgroups of ASD. However, one argument against the hypothesis of specific developmental delay is that some skills that develop soon after birth in typically developing children, such as eye contact, may continue to be absent in ASD, even when individuals pass ToM tasks. A possible explanation for this seemingly contradictory finding is that, as ASD individuals who pass false belief reasoning tasks require much greater chronological and verbal mental ages than typically developing children, these children might solve ToM tasks using language skills and general cognitive processes rather than specific mental state processes, or innate ToM skills (Bauminger & Kasari, 1999; Tager-Flusberg, 2007). Individuals with ASD activate different brain regions to control subjects when answering ToM questions (Baron-Cohen et al., 2008) and take longer to answer ToM questions than typically developing individuals, which points to reasoning out the answer rather than understanding emotions and beliefs (Kaland, Smith, & Mortensen, 2007). These results suggest that compensatory skills may increasingly be used by older children with ASD. However, there is currently little evidence on which to base developmental theories, and much of the research suffers from the fact that either only certain subgroups of ASD are included, or that performance is not differentiated by subgroup (see Table 2 for a summary of studies examining ToM development in ASD). Hence, it is unclear whether ToM skills are similar throughout ASD or whether certain clusters of ToM ability are present. The profile of ToM skills in the different ASD subgroups will be discussed next.

Table 2.

Summary of the Studies Investigating Development of Theory of Mind in Autism Spectrum Disorders

Study	Groups	Study Type	*N*	Age range	Tests	Results
Holroyd & Baron-Cohen, 1994	LFA	7-8 year follow-up	17	6-16[a]	1st and 2nd order FB	No significant difference in ToM scores
Ozonoff & McEvoy, 1994	HFA, PDD-NOS	3 year follow-up	34	10-23[a]	1st and 2nd order FB, Double Bluff	No significant difference in ToM scores
Paynter & Peterson, 2009	HFA, AS, TD	Cross-sectional	63	4 - 13	1st and 2nd order FB	Significant increases in ToM with age. ToM ability predicted by syntactic ability.
Steele, Joseph & Tager-Flusberg, 2003	Autism[b]	1 year follow-up	57	4 – 14[a]	Desire, Pretend Play, Perception-Knowledge, Sticker Hiding, 1st and 2nd order FB, Lies and Jokes, Traits, Moral Judgement	Significant increases in ToM.

Note. Only studies that compared children at different ages were included in this table.
ToM = theory of mind, FB = false belief, HFA = high-functioning autism, AS = Asperger's syndrome, PDD-NOS = pervasive developmental disorder not otherwise specified, TD = typical development

[a] Age at first assessment

[b] This group included both low-functioning and high-functioning children

Differences in ToM within the autism spectrum. Thus far, ToM has been described within ASD in general. However, as varying degrees of symptomatology, verbal ability and intellectual disability have been described between the different subgroups, it follows that level of ToM may also differ within subgroups. Of particular importance is whether differences in ToM exist even when cognitive ability is accounted for; this would suggest fundamental taxonomic differences in groups' 'autisticness' and would therefore support the idea of different diagnostic groups within autism. In the long run, understanding the fundamental differences between ASD subgroups, if indeed there are any, may provide clues as to whether the ASDs are due to similar or different causes.

High-functioning and low-functioning autism. Most of the research has focused on these groups, particularly the high-functioning autism group, as participants from this group are easy to obtain and test. As has been noted, a certain level of verbal ability seems to be required to be able to pass first-order false belief tasks. Therefore, as children with high-functioning autism have higher verbal ability scores than children with low-functioning autism, this group generally performs better on ToM tests (Dyck, Ferguson, & Shochet, 2001).

High-functioning autism and Asperger's syndrome. Descriptions of individuals with ASD often highlight differences in social behaviour between Asperger's syndrome and high-functioning autism. Individuals with Asperger's syndrome have been described as having an 'active but odd' personality style, in comparison with most children with autism who are 'passive and aloof' (Wing, 1981). Unlike individuals with autism, they often desire friendships and romantic relationships, and show more advanced social reactions, for example in greeting, showing affection and taking pleasure in social interactions (Frith, 2004; Ghaziuddin, 2008). Given the importance of social communication, both in daily life and in the diagnosis of ASD, remarkably few studies have compared high-functioning autism and Asperger's syndrome individuals on formal tests of ToM.

The vast majority of studies investigating ToM in ASD without concurrent intellectual disability make no distinction between high-functioning autism and Asperger's syndrome. This trend was followed earlier in the literature review with the discussion of delay and deviance in ToM because of this paucity in the literature. However, an examination of the similarities and differences in ToM between high-functioning autism and Asperger's syndrome is warranted here, particularly in the light of the proposed changes to the DSM.

To my knowledge, only eight studies have compared high-functioning autism and Asperger's syndrome, three of which have been published within the last two years, when the proposed changes to the DSM were announced. Of the eight, four studies have found differences in ToM (Kuroda et al., 2011; Ozonoff, Rogers, & Pennington, 1991; Paynter & Peterson, 2009; Ziatas, Durkin, & Pratt, 1998) - with the Asperger’s syndrome group performing better than the high-functioning autism group - and four studies found no differences in ToM ability (Dahlgren & Trillingsgaard, 1996; Dissanayake & Macintosh, 2003; Klin, 2000; Spek, Scholte, & Berckelaer-Onnes, 2009). It should be noted however that Dissanayake and Macintosh (2003) *did* find differences between the groups on first-order false belief tasks, but when using only the children who passed these tasks, did not find differences between the groups on second-order false belief. Similarly, Klin (2000) used only children who had already passed second-order false belief tasks in his study. If ToM ability is better developed in Asperger’s syndrome than in high-functioning autism, as some of the studies suggest, this procedure may have unintentionally biased the selection so that the high-functioning autism individuals participating had greater theory of mind skills than is normally found within this group.

Another difficulty in comparing the studies is that different tests of ToM and different diagnostic criteria for Asperger’s syndrome were used. Some studies used the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 2001) to aid in diagnosis, other simply report that diagnosis was done by a psychologist or psychiatrist according to DSM criteria, and Ozonoff et al. (1991) allowed children with Asperger's syndrome who had language delays, and thus according to the DSM-IV criteria should be diagnosed with autism or PDD-NOS. Of the eight studies, two used first-order false belief and/or basic emotion attribution tasks (Paynter & Peterson, 2009; Ziatas et al., 1998), three used first- and second order false belief tasks (Dahlgren & Trillingsgaard, 1996; Dissanayake & Macintosh, 2003; Ozonoff et al., 1991) and three used advanced theory of mind tasks, including Strange Stories, interpreting a television clip, and attributing emotions to geometric shapes (Klin, 2000; Kuroda et al., 2011; Spek et al., 2009). The details of tests and diagnostic procedures used can be found in Table 2. Integrating these results, the trend seems to be that individuals with Asperger’s syndrome perform better than individuals with high-functioning autism on basic ToM tasks, but not on more advanced ToM tasks (but see Kuroda et al., 2011). However, there is clearly still a need to compare high-functioning autism and Asperger’s syndrome on a variety of ToM tests of different difficulty levels in order to interpret previous contradictory results.

Table 3.

Summary of the Studies Comparing Theory of Mind in High-Functioning Autism and Asperger's Syndrome

Study	Sample Size	Age range	Tests	ToM Results
Dahlgren & Trillingsgaard, 1996	HFA, AS = 20	6 – 16	1st and 2nd order FB	AS = HFA < TD
Dissanayake & Macintosh, 2003[a]	HFA = 21, AS = 19, TD = 20	5 - 11	1st and 2nd order FB Social Behaviour	1st order FB: HFA < AS = TD 2nd order FB: HFA = AS = TD Social Behaviour: AS = HFA < TD
Klin, 2000[a]	HFA, AS, TD = 20	HFA: 20.5 (10.8) AS: 18.9 (11.8)[b]	Social Attribution Task	AS = HFA < TD*\
Kuroda et al., 2011	HFA = 17, Other PDD (AS & PDD-NOS) = 11, TD = 50	16 - 45	Describe TV character's mental state from visual or auditory information	HFA < AS = TD
Ozonoff, Rogers & Pennington, 1991	HFA = 13, AS = 10	8 – 21	1st and 2nd order FB	HFA < AS = TD
Paynter & Peterson, 2009	HFA = 19, AS = 24, TD = 20	5 - 12	1st order FB, Belief-Emotion	HFA < AS = TD
Spek, Scholte & Berckelaer-Onnes, 2009	HFA = 32, AS = 29, TD = 32	18 - 60	Strange Stories Faux Pas Reading the Mind in the Eyes Empathy Quotient	AS = HFA < TD
Ziatas, Durkin & Pratt, 1998	HFA, AS, TD, SLI = 12	HFA: 8.5 (2.7) AS: 6.11 (2.1)[b]	1st order FB Belief Term Comprehension & Expression	HFA < AS = TD = SLI

Note. ToM = theory of mind, FB = false belief, HFA = high-functioning autism, AS = Asperger's syndrome, TD = typical development, SLI = specific learning impairment

[a] Denotes studies that have used only children who passed false belief tasks

[b] Age ranges were not reported for these studies; means and standard deviations are given

ToM in PDD-NOS. Very little research has been done on ToM in PDD-NOS. There is some evidence that children with PDD-NOS perform equally well on basic emotion recognition tasks (Serra, Jackson, van Geert, & Minderaa, 1998) and on emotion role-taking tasks (Serra et al., 1995) as IQ-matched controls. However, most studies that have compared PDD-NOS and typically developing children suggest that, like children with autism, children with PDD-NOS are impaired on various ToM and emotion-recognition tasks (Buitelaar, van der Wees, Swaab–Barneveld, & van der Gaag, 1999; Serra et al., 2002; Sicotte & Stemberger, 1999).

As far as I am aware, only two groups have compared ToM performance in PDD-NOS with that in autism; that of Begeer and colleagues and Buitelaar and colleagues. Buitelaar and colleagues found no difference between the PDD-NOS and autism groups on emotion-recognition, first or second-order false belief (Buitelaar et al., 1999; Buitelaar, van der Wees, Swaab-Barneveld, & van der Gaag, 1999). Begeer and colleagues found that children with autism and children with PDD-NOS performed similarly poorly on correcting an examiner's false belief, but that the PDD-NOS group's performance increased when promised a reward, whereas the autism group's performance did not change (Begeer, Rieffe, Meerum Terwogt, & Stockmann, 2003). These results suggest that ToM deficits might be milder in PDD-NOS than in autism, but more research is necessary to confirm this, especially since the studies conducted thus far have only used children with high-functioning PDD-NOS and autism (barring Sicotte & Stemberger, 1999). This means that previous results may not apply to the groups as a whole.

To summarise, numerous studies have found that children with ASD show severe deficits in ToM compared with both typically developing and intellectually disabled children, which throws doubt on the notion that ToM is based on some uniform cognitive substrate or ability. Within ASD, the pattern of ToM performance is less clear. Children with high-functioning autism seem to outperform children with low-functioning autism on ToM tasks, possibly because of the poorer verbal skills found in the latter group. Few studies have compared the relatively new DSM-IV diagnostic categories of PDD-NOS and Asperger's syndrome to autism. Interpretation of study results are also obstructed by diagnostic criteria that have been vaguely defined or that differ between studies. Hence, much controversy still exists over whether ToM ability is equivalent in high-functioning autism and Asperger's syndrome. Research suggests that individuals with Asperger's syndrome perform better than individuals with high-functioning autism on basic ToM tasks, but not on more advanced ToM tasks. PDD-NOS cases seem to have the same impairment in ToM as is found in autism, though these deficits may be less severe. Because of the paucity of information in this area it is important not only to compare these diagnostic categories before joining them into a single category, but also – specifically because of diagnostic difficulties that exist - to look at naturally occurring groups of ToM skills or disability, regardless of diagnostic category.

With regard to development, some evidence has been presented that ToM skills may develop in ASD. However, results from the initial studies are contradictory as to whether the developmental sequence for ToM is the same in autism as in typical development and whether ToM development may plateau in ASD. Clearly more research is needed to understand how stable ToM is across the lifespan in individuals with ASD.

In light of changes to the DSM classification from five diagnostic categories to one autistic disorder category, a new dimensional classification method will become necessary to distinguish between individuals at different levels of severity of the disorder. ToM ability might be one way to identify such homogenous subgroups within autistic disorder: it reflects some of the core problems of socialization and communication in ASD, which are associated with the adaptive functioning of individuals. Thus, this study aims to establish whether ASD children of different ages and abilities can be differentiated into relatively distinct groups of ToM ability. This will be done by looking at how ToM skills in ASD change with age, and by endeavouring to identify clusters of ToM ability within ASD. These clusters may correspond to current DSM diagnostic categories, or may suggest new dimensional classifications in ASD. In short, this study examined whether any obvious patterns of ToM ability exist, both cross-sectionally and in terms of development.

Rationale for Research

Stability of ToM

It is still unclear to what extent ToM development takes place in children with ASD, and whether ToM development is prevalent and similar throughout the spectrum. As far as I am aware, no study has investigated ToM development in all the different subgroups of ASD. The results of previous studies on ToM development have also been contradictory: recent studies with large sample sizes and, in particular, young participants have found significant increases in ToM, while longitudinal studies with long follow-up periods have found no differences in ToM over time. Clearly, more research is needed using a wide variety of ToM tests to be able to identify development. Specifically, a large improvement in ToM is seen during the preschool years in typical development. I therefore aimed to include young, preschool-aged participants in my ASD sample. The following questions regarding ToM development were asked:

(1) Do ToM abilities develop in children with ASD?

(2) If ToM development does take place, is this development limited to high-functioning autistic and Asperger's syndrome children, or does it occur in all ASD participants?

(3) Furthermore, if ToM development does take place in children with ASD, how does it compare with ToM development in typically developing children in terms of

- (a) rate of development,
- (b) age of onset of early-developing ToM abilities
- (c) highest level of ToM obtained (i.e., is there an observable plateau in ToM development in children with ASD)?

ToM Universality and its Profile in the ASD Subgroups

In light of the proposed decision to merge the DSM –IV categories autism, Asperger's syndrome and PDD-NOS into a single autistic disorder category, I aimed to establish whether ASD children of different ages and abilities could be differentiated into relatively distinct groups of ToM ability. If no clusters of ToM ability emerge after confounding cognitive variables are factored out, this lends credence to the idea of a singular autistic disorder category. If clusters similar to the present DSM diagnostic categories emerge, it should give us pause before combining the different diagnostic categories into a singular category. Lastly, if clusters of ToM ability emerge, but these do not correspond to the current diagnostic categories, it indicates that a new diagnostic classification may need to be devised.

Particular attention is focussed on the controversial DSM categories of Asperger's syndrome and PDD-NOS. I examined whether there are observable differences in ToM between high-functioning autism and Asperger's syndrome and whether a homogenous ToM profile exists for the category PDD-NOS.

Methods

Research Design

The study was an existing groups-comparison and was divided into two sections: a cross-sectional comparison of five groups; low-functioning autism, high-functioning autism, Asperger's syndrome, PDD-NOS, and a typically developing control group. A cross-sectional study can show that older children have greater ToM skills than younger children, but cannot prove beyond doubt that development took place as different groups of children are tested. A group of children were therefore reassessed on the ToM tasks after 1 year. Because of time constraints, a smaller sample of 10 children was chosen.

Participants

Participants were recruited from the Western Cape, Kwa-Zulu Natal and Gauteng provinces. Convenience and snowball sampling were used to obtain participants for the study: firstly, autism-specific, special needs and mainstream schools were approached to obtain participants; secondly, autism support groups were contacted; and thirdly existing participants were asked to identify potential suitable candidates for the study.

All ASD participants were diagnosed by qualified clinicians independent of this study. Participants with a diagnosis of autism were further divided into the categories of low-functioning or high-functioning autism based on their IQ. Autistic participants with an IQ of 70 or below were in the low-functioning autism group, while participants with an IQ of higher than 70 were in the high-functioning autism group.

All participants were between the ages of 4 and 16 years. This age range was chosen as false belief reasoning should be present by 4 years of age. As the most difficult tasks were appropriate for typically developing 11-year-old children, the cut-off age of 16 years was chosen so that if delayed development does take place in ASD, the oldest children might be able to pass the advanced ToM tasks. The 10 children chosen for follow-up were all diagnosed with high-functioning autism and were between the ages of 5 and 10 years. I chose this group as this age range was most likely to show development. A high-functioning autism group was also the easiest to access and test, and previous research suggested that this group was likely to show improvements in ToM (e.g., Steele et al., 2003).

Inclusion and exclusion criteria. To be included in the ASD group, participants had to have been diagnosed with ASD by a clinician independent of this study. Children with a history of head injury or infantile meningitis were excluded from the study. ASD candidates with additional neurological conditions were excluded from the study, as were control candidates with any neurological conditions. Individuals with any serious social deficits, such as conduct disorder (CD) or oppositional defiant disorder (ODD), a communication disorder, attention-deficit hyperactivity disorder (ADHD) or any pervasive developmental disorder, or a history of these disorders, were excluded from the typically developing control group. This information was obtained through the demographic questionnaire (see Appendix B). A further condition was that all participants had to be fluent in English or Afrikaans.

Ethical considerations. This study followed the ethical guidelines for research with human subjects outlined by the Health Professions Council of South Africa (HPCSA) and the University of Cape Town (UCT) Codes for Research. Permission was obtained from the relevant provinces' Education Departments and schools to approach their students to participate in the study (Appendix C).

Written informed consent was obtained from the participants' parents or guardians beforehand, and informed assent was obtained from the participants on the day of testing (see Appendix D). Participants were assured that data would be kept confidential, and would only be used for research purposes. This study involved minimal risk and no potentially harmful tasks were involved. The only possible risk was that children could become uncomfortable or fatigued during testing. Children were allowed to take a break between tasks if they became tired and could withdraw from the study with no negative consequences if they no longer wished to take part. Parents were given written feedback regarding their child's performance after the study.

Initial sample. One hundred and ninety seven children were volunteered to take part in the study, of which 162 had ASD. In the ASD group, 29 participants were excluded immediately because they were non-verbal, 17 because they had additional neurological disorders besides ASD (meningitis, Landau-Kleffner syndrome, ataxic cerebral palsy, vision problems), two because of a previous head trauma, and one because of an incorrect diagnosis. In the typically developing group, three children were excluded because of a diagnosis of attention deficit/ hyperactivity disorder and one because of a previous language delay. 143 children therefore participated in the research.

The ASD participants were matched as closely as possible with typically developing children on age, gender, socio-economic status and home language.

Procedure

Testing took place at the participant's school or home, or if such an arrangement could not be made, at the Department of Psychology. Participants were tested individually in a quiet room free of distractions. It was planned that testing would take place over two sessions of approximately 45 minutes each for children younger than 6 years, and 80 minutes each for children 6 years of age and older. However, in the case of preschool children and children with ASD – specifically the low-functioning group – more and shorter sessions were often used.

During the first session, the Comprehension of Instructions task was administered. Children had to correctly complete at least two-stage verbal instructions on this task to be included in the rest of the study. Next, tests of general intellectual functioning, working memory and processing speed were administered. In the second session, the ToM tasks and the executive functioning tasks Verbal Fluency and Colour-Word Interference were administered. During the assessment of ToM, printed stories with accompanying pictures were left in front of the child to minimize memory and linguistic demands. Details on the tests are given below.

The 10 children who were identified for follow-up were contacted again one year later and were tested once more on the ToM and general cognitive measures.

Measures

A variety of tasks assessing general cognitive functioning and ToM skills were used for the research. These tests were not developed or normed in South Africa. However, this should not affect the interpretation of the results, as the control group provided a benchmark of what a typically developing South African child's performance should be like.

Comprehension of instructions. The Comprehension of Instructions task from the NEPSY-II (Schmitt & Wodrich, 2004) was administered to all ASD participants. The Comprehension of Instructions task was developed for children aged 3-16 years and assesses the ability to receive, process, and execute oral instructions of increasing syntactic complexity.

Cognitive ability.

General intellectual functioning. To assess general intellectual functioning, the *Wechsler Abbreviated Scale of Intelligence* (WASI; Wechsler, 1999) was used for participants aged 6 years and older, and the third edition of the *Wechsler Preschool and Primary Scale of Intelligence* (WPPSI-III; Wechsler, 2002) for participants younger than 6 years. All four subtests of the WASI were administered to obtain verbal, performance and full scale IQ scores. All the subtests of the WPPSI-III for 2-3 year olds were administered. A short form of the WPPSI-III for 4-7 year olds was administered to obtain verbal and full scale IQ scores (Sattler, 2008). For the 4-7 year olds' version of the WPPSI-III, the Vocabulary, Information and Similarities subtests were administered to obtain verbal IQ. Full scale IQ was calculated from participants' performances on the Information, Matrix Reasoning, Picture Completion, Symbol Search and Similarities subtests.

Three subtests from the fourth edition of the *Wechsler Intelligence Scale for Children* (WISC-IV UK; Wechsler, 2004) were administered to obtain a measure of working memory and processing speed. Performance on the Digit Span task was used as a measure of working memory. To obtain a measure of processing speed, the Coding and Symbol Search subtests were used.

Executive functioning. The *Delis-Kaplan Executive Function System* (D-KEFS; Delis, Kaplan, & Kramer, 2001) measures key components of executive function. The Verbal Fluency and Colour-Word Interference subtests from this battery were administered. Verbal Fluency was used to measure both lexical and semantic generativity. The Colour-Word Interference task, which is based on the Stroop (1935/1992) test, was used to measure a participant's ability to inhibit an overlearned verbal response.

Theory of mind. The ToM battery consisted of 11 ToM tasks that were divided into four modules of increasing difficulty. The ToM battery was adapted from that used by Steele et al. (2003), and used similar tests in the early and basic modules, but differed from these authors' battery in that the Perception-Knowledge task was moved from the basic to the early module in this study, as it was judged that this ability develops somewhat earlier than false belief reasoning. The Explanation of Action task was added to the basic module and two of the advanced tasks (namely, Traits and Moral Responsibility) were replaced with other tasks (Strange Stories and Faux Pas) that measure ToM ability directly, rather than performance on factors associated with ToM. Because of the large jump in difficulty level between the basic and advanced modules, the battery was divided into four, rather than three, difficulty levels. This decision was taken so that children would not have to complete tests that were inappropriately difficult for their chronological or mental age, and to prevent children from becoming demotivated and unwilling to complete the assessment.

The tasks included both test and control questions, with the exception of the Pretend and Sticker Hiding tasks. All tasks, except those using dolls and the advanced Faux Pas task, had accompanying pictures to minimize linguistic and memory demands. Participants' scores on the various ToM tasks were added to obtain a Total ToM score. All the tasks were also scored as pass or fail.

Early module. For each of the tasks on the early module, participants had to get at least 3 out of 4 questions right to obtain a pass. The Pretend Play task was designed for this study and tested the ability to use a doll as an independent agent in a pretend situation. The original task by Kavanaugh, Eizenman and Harris (1997) and Steele et al. (2003) was judged to be inappropriately female-gender stereotyped. In light of the fact that most children with autism are boys, the stories were changed to depict gender-neutral events. The task consisted of four structured play scenarios. Participants were asked to complete scenarios by acting out each event using the dolls (see Appendix E).

The Desire task (Steele et al., 2003; Wellman & Woolley, 1990) tested the ability to predict behaviour based on a character's stated desire. The task contained two picture stories. In each of the two stories the protagonist looked for an object that was in one of two named locations. The character failed to find the object in the first location. Participants were asked what the character would do next and why.

The Perception-Knowledge task (Pratt & Bryant, 1990; Steele et al., 2003) tested the ability to know that a character obtains knowledge from visual access. The task contained four control questions, alternated with four test questions in which one doll looked into a box and another pushed the box. The child was asked which doll knows what is inside the box.

Basic module. The Location-Change False Belief task (Baron-Cohen et al., 1985; Wimmer & Perner, 1983) contained two picture stories wherein an object is moved while the main character is out of the room. The child was asked whether the character knows where the object is, where the character would look for the object, and why the character would look there. To obtain a pass on this task, participants had to get at least 6 out of 9 questions right.

In the Unexpected-Contents False Belief task (Perner, Leekam, & Wimmer, 1987) participants were shown four familiar containers with unexpected contents inside. The child was asked, "When you first saw the box, all closed up, what did you think was inside?" and

"When X comes into the room, and sees the box all closed up, what will he/she think is inside?" To obtain a pass on this task, participants had to get a score of at least 8 out of 12.

The Explanation of Action task (H. B. Tager-Flusberg, personal communication, March 14, 2008) contained 12 stories in which an action based on an emotion, desire, cognitive process (think, know or forget) or non-mental event was portrayed. The child was asked why the character performed the action. If the child was unable to answer, he or she was prompted with "what is going on in [the character]'s head when [the character performed the action]?" To obtain a pass on this task, participants had to get a score of at least 8 out of 12.

The Sticker Hiding task (Devries, 1970) is a deception task wherein the participant is required to hide a sticker from the experimenter. This task was included as a naturalistic and non-verbal ToM test as it required the participant to take another person's perspective and knowledge into account in order to keep the location of the sticker a secret (for example, while the sticker would be hidden if only the hand containing the sticker were closed, it would be immediately obvious to an observer where the sticker is). The task started with six practice trials wherein the experimenter hid a sticker in one, both or neither hand so that the child would guess the location of the sticker correctly at least once, and incorrectly at least once. The child then hid the sticker for ten trials, of which the last five were scored for ability to hide the sticker from the experimenter. Points were given for (1) taking both hands behind the back to hide the sticker, (2) bringing both hands to the front, (3) keeping both hands closed until the experimenter has made a guess, (4) keeping the sticker completely invisible in the hand and (5) using a switching strategy for both guessing and hiding. To obtain a pass on this task, participants had to obtain a score of at least 16 out of 22.

Intermediate module. The Second-Order False Belief task (Ozonoff & McEvoy, 1994) consisted of two picture stories, and tested knowledge of a character's beliefs about a second character's beliefs. The child was asked an ignorance ("Does Mom know what you are making her for Mother's Day?"), belief (What does Mom think you are making her for Mother's Day?"), and justification question. To obtain a pass on this task, participants had to get a score of at least 6 out of 8.

Strange Stories (Happé, 1994) consisted of 18 illustrated stories of 13 types: non-mental inference (control), lie, white lie, joke, pretend, double-bluff, persuasion, forgetting, misunderstanding, figure of speech, appearance-reality, irony and contrary emotions. Participants were asked whether what the character said was true, and why the character said it. To obtain a pass on this task, participants had to get a score of at least 26 out of 36.

Advanced module. The Lies and Jokes task (Steele et al., 2003; Winner, Brownell, Happé, Blum, & Pincus, 1998) consisted of four picture stories, two of which contained a lie and two a joke. In each story, a child character said something that their parent knew to be untrue. In the joke version, the child character knew that the parent knew the truth, while in the lie version the child did not know that the parent knew the truth. Participants were asked whether the child's statement is a lie or a joke. To obtain a pass on this task, participants had to get a score of at least 12 out of 16.

The children's version of the Faux Pas task (Baron-Cohen et al., 1999; Stone, Baron-Cohen, & Knight, 1998) was administered. This task contained five control stories which depicted a normal social event, and five test stories wherein a character said something awkward or embarrassing. After reading the story, participants were asked, "Did anyone say something they shouldn't have said or something awkward?" If they responded yes, they were asked, "Who said something they shouldn't have said or something awkward?", "Why shouldn't he/she have said it or why was it awkward?", and "Why do you think he/she said it?". The original task by Baron-Cohen and colleagues (1999) contained ten control and ten test questions. The test and control questions are the same in all respects, except for one detail that changes (a character either makes a neutral statement or makes an uncomfortable statement). However, it was thought that this repetition of stories could be confusing for children, causing them not to concentrate on the second story, or to give the same answers for both stories, which may not be correct for both versions. Therefore, the test was split into two versions (each containing five control and five test stories), so that if a certain test question was in Version A, its matching control was in Version B, and thus stories were not repeated. I alternated between the versions. To obtain a pass on this task, participants had to get a score of at least 28 out of 40.

To keep the duration of the assessment short, but obtain a reliable measure of ToM ability, children started the ToM battery at the module most appropriate for their age or developmental level and continued until failing a module or reaching the end. Therefore, typically developing children between the ages of 3 and 5 years who, according to the international literature, should pass tasks like Pretend Play and Desire, started on the early ToM module. All children with autism and PDD-NOS also started on the early ToM module, and proceeded to the next module if they attained at least half the maximum score. Typically developing children between the age of 6-7 years and children with Asperger's syndrome started on the basic module, as it was judged that they should comfortably pass false belief tasks. Typically developing children 8 years and older started on the intermediate module. The skills tested in this module, namely second-order false belief and understanding non-literal speech, should be present in typical development by 6-7 years old. If participants who started on the basic and intermediate modules attained at least half of the maximum score, those children received full credit for the easier modules that they did not do and advanced to the next module until they reached the end or attained less than half of the maximum score. If children attained less than half the maximum score on their starting battery, they were tested on the previous battery. Wellman and Liu (2004) found that if children passed a more advanced ToM test, they passed all easier ToM tests. Thus, the validity of the results should not be affected by children not completing all the modules.

All the aforementioned tasks were piloted in 2008 on typically developing and high-functioning autistic children. The ToM battery was found to be appropriate for a South African context. The results of the pilot study are given in Appendix F. After the pilot study, some questions on the intermediate and advanced modules were adjusted to reduce language demands and be more culturally appropriate. Specifically, the Lies and Jokes task did not show a clear performance increase with age during the pilot study. It is possible that participants always interpreted the child character’s statement, “I did a really good job [on the task]”, as a lie rather than as a joke because a joke seemed inappropriate in a setting where the child is talking to an authority figure. The phrasing of the Lie/Joke questions on this test was therefore changed to that in the adult version of this task. Instead of asking “Was [the character] lying or just joking?”, participants were asked “Was [the character] lying so he/she wouldn’t get caught, or joking because he/she was embarrassed?” I hoped that by including the speaker’s motivation for the false utterance the distinction between lying and joking would be made clear, so that participants would not fail due to different cultural views of what is acceptable within a parent-child interaction, or different interpretations of the words *lying* and *joking*. Some of the words on the Strange Stories task were changed to similar words that would be known by South African children, for example “John was in one of the *cubicles* in the bathroom at school” was changed to “John was in one of the *toilets* in the bathroom at school”. I also found that the meaning of the word *awkward* (used in the Faux Pas task) needed to be explained to participants beforehand.

Data Analysis

All statistical analyses were completed using PASW Statistics version 18.0 (SPSS, 2010) and Statistica version 9 (StatSoft, 2010). To explore the development of ToM ability, five separate regressions were performed (typical development, low-functioning autism, high-functioning autism, Asperger's syndrome, PDD-NOS) with ToM as the dependent variable and Age as the independent variable. The correlation coefficients, slopes and intercepts of these regressions were compared using GraphPad Prism 5 (Motulsky, 1999). A *t*-test was used to compare ToM scores at initial assessment and follow-up 1 year later.

To compare group performances on the ToM battery, a one-way analysis of variance was done on ToM score with Group (low-functioning autism, high-functioning autism, Asperger's syndrome, PDD-NOS, and typical development) as the independent variable. An analysis of covariance was also conducted with age, verbal IQ, inhibition/set shifting, verbal generativity, digit span and processing speed as the covariates. The assumption for homogeneity of variance was not upheld, but the residuals were normally distributed, so the results of the analysis remain reliable. A *K*-means cluster analysis (non-hierarchical clustering) was used to group together cases by calculating the Euclidean distances between cluster centroids for performance on the various ToM tasks. G*Power version 3 (Faul, Erdfelder, Buchner, & A. Lang, 2009) was used to calculate effect sizes and achieved power.

Results

Sample Characteristics

Of the 113 ASD children who participated in the research, a further 21 were excluded because they could not correctly complete at least two-stage verbal instructions (“show me the big, happy bunny”) on the NEPSY-II Comprehension of Instructions task (Schmitt & Wodrich, 2004), three children were too anxious to complete the tasks and two children did not complete the assessment. Only two children were excluded from the typically developing group; both because they did not wish to complete the tasks. The final sample therefore consisted of 86 children with ASD (21 low-functioning autism, 24 high-functioning autism, 21 Asperger's syndrome and 20 PDD-NOS) and 30 typically developing children between the ages of 4 and 16 years.

The groups were matched on as closely as possible on gender, age, language and socio-economic status. Socio-economic status distribution within the sample was determined from annual household income, parental education level and number of household assets, as described in Myer, Stein, Grimsrud, Seedat, and Williams (2008). The list of household assets included items such as having access to running water, owning a fridge, having a credit card, and so forth.

Language. There were significant differences in the distribution of home language between the groups. As the typically developing group was most representative of the language distribution in the Western Cape (Statistics South Africa, 2006), with 27% Afrikaans, 23% Xhosa, 47% English and 3% other home language speakers (compared with an estimated 41% Afrikaans, 29% Xhosa, 28% English and 2% other home language speakers in the Western Cape), the group differences found may be an indication that many children with ASD are either unidentified or misdiagnosed in certain language groups. This may be particularly true of the Asperger’s syndrome and PDD-NOS groups, which were almost exclusively white and English speaking. The potential racial bias in diagnosis that was observed mirrors previous results from the US, where children with Asperger’s syndrome and PDD-NOS are significantly more likely to be white and less likely to be Hispanic (Rosenberg, Daniels, J. K. Law, P. A. Law, & Kaufmann, 2009).

General intellectual functioning. Although children in the high-functioning autism and Asperger’s syndrome groups all had IQ scores in the non-intellectually disabled range (bar one child diagnosed with Asperger's syndrome who had an IQ of 67), there was still a significant difference in full scale IQ between these groups and the typically developing group (p < .05). The high-functioning autism group had a lower average IQ score than the typically developing group, while the Asperger's syndrome group on average had a higher IQ than the typically developing group. The PDD-NOS group had the largest range of IQ scores, and included children with and without intellectual disability. All the children in the low-functioning autism group had IQ scores of 70 and below, in the mild to moderate intellectual disability range.

Age. The Asperger's syndrome group was slightly older than the other groups, though not significantly so. The older average age in this group reflects the older age at diagnosis for children with Asperger's syndrome. Howlin and Asgharian (1999) reported that in the UK the average age that parents seek medical help for children later diagnosed with Asperger's syndrome is 3.49 years and that, on average, the diagnosis is only confirmed at 11.13 years. The demographic characteristics for the sample are shown in Table 4.

Table 4.

Demographic Information for the Sample

	LFA	HFA	AS	PDD-NOS	TD	F/χ^2	Significance
N	21	24	21	20	30		
Age						2.06	.091
Mean	10.37	8.91	11.12	9.68	9.28		
SD	2.94	2.89	2.38	2.68	3.43		
Range	5-16	5-16	7-15	5-13	4-16		
Sex						4.54	.338
Male	17	22	20	17	23		
Female	4	2	1	3	7		
Home Language						29.79	.003[a]
English	15	17	19	20	14		
Afrikaans	2	3	1	0	8		
Xhosa	3	2	0	0	7		
Other	1	2	1	0	1		
SES						9.59	.295
Low	6	13	3	7	9		
Medium	5	5	8	7	12		
High	6	5	9	5	7		
Did not disclose	4	1	1	1	2		
FSIQ						32.40	< .001
Mean	61.52	79.62	99.38	80.75	90.21		
SD	6.58	8.19	13.55	11.34	14.73		
Range	50-70	71-106	67-123	66-103	62-122		

Note. SES = socio-economic status, FSIQ = full scale IQ, LFA = low-functioning autism, HFA = high-functioning autism, AS = Asperger's syndrome, PDD-NOS = pervasive developmental disorder not otherwise specified, TD = typical development

[a]However, this result cannot be meaningfully interpreted because of the unreliability of Chi-squared tests in the presence of empty cells.

ToM Development

Typical development. The scatterplot for the typically developing group with age and ToM as the independent and dependent variables respectively is shown below. Theory of mind seems to develop rapidly within the first 9 years, after which development slows. Because the scatterplot seemed to indicate a nonlinear relationship, the runs test for nonlinearity (Motulsky, 1999) was conducted for all 5 groups. None of the regressions significantly deviated from a linear Age-ToM relationship however (all $p > .05$), indicating that the data were appropriate for a linear regression analysis.

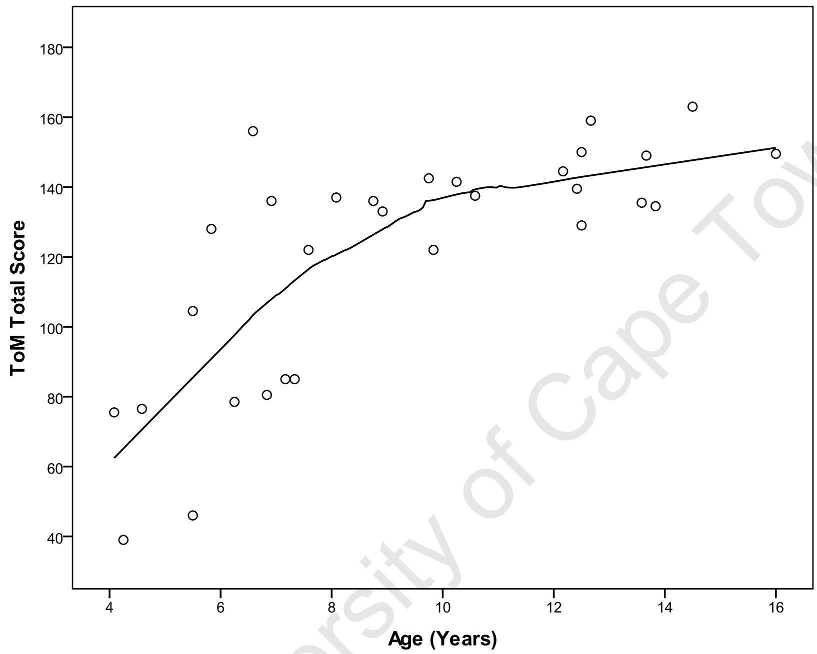

Figure 1. The correlation between theory of mind and age in typical development. Theory of mind seems to develop rapidly within the first 9 years, after which development slows.

Delayed versus deviant development in ASD. The development of ToM ability was examined (1) by looking at the performance on sequentially more difficult ToM tasks, (2) by comparing the time of onset of ToM abilities, as measured by the battery, in the ASD subgroups and (3) by comparing the different developmental rates of ToM in ASD.

On average, all the groups did the best on the Pretend Play task and the worst on the Faux Pas task (see Table 5). Although there were small differences in the percentage of children within a group to pass the tasks, there were no large differences in the order in which the tasks were passed, which seems to indicate that ToM skills develop in the same order in ASD as in typical development.

Table 5.

Percentage of Participants Within Each Group Who Passed the Theory of Mind Tasks

	LFA	HFA	AS	PDD-NOS	TD
Early Module					
Pretend Play	65	92	100	100	100
Perception-Knowledge	10	52	91	70	100
Desire	25	56	91	65	90
Basic Module					
Sticker Hiding	20	68	91	65	100
Explanation of Action	30	40	100	70	90
Unexpected-Contents False Belief	15	32	91	65	93
Location-Change False Belief	5	12	91	65	90
Intermediate Module					
Strange Stories	0	4	62	10	50
2nd Order False Belief	0	0	48	15	40
Advanced Module					
Lies and Jokes	-	0	48	15	40
Faux Pas	-	4	43	15	30

Note. Passing criteria were set so that a pass indicated above-chance performance on the task. For specific passing criteria, see the Methods section.
LFA = low-functioning autism, HFA = high-functioning autism, AS = Asperger’s syndrome, PDD-NOS = pervasive developmental disorder not otherwise specified, TD = typical development

In the typically developing group all of the children passed the tasks in the early module except for two children (one 4 and one 5 years old) who failed the Desire task. Sixty seven percent of 4-6 years olds and all children 7 years and older passed the false belief tasks. On Second-Order False Belief, most children passed from 8 years onwards. On Strange Stories, all children under 5 years failed, 50% of children between 5 and 10 years passed, and almost all children older than 10 years passed (only one older child failed). On Lies and Jokes and Faux Pas all children under the age of 8 years failed; the majority (75%) passed Lies and Jokes at 10 years and most children older than 12 years passed Faux Pas. The age ranges at which different ToM skills were present in the typically developing South African sample were therefore roughly consistent with previously reported international results (Brüne & Brüne-Cohrs, 2006; Sullivan et al., 1994; Wellman & Woolley, 1990).

However, Table 5 shows that some ToM tasks were more difficult than other tasks in the same module. For instance, all the groups found the Desire task much more difficult than the other tasks in the early module, and more difficult even than some of the basic ToM tasks such as Sticker Hiding and Explanation of Action. Many of the children also found the Second-Order False Belief task more difficult than Strange Stories, though previous research suggested that second-order false belief is necessary to understand the difference between many non-literal statements, such as the difference between a lie and an ironic statement. Interestingly, the low and high-functioning autism groups performed much better on the Unexpected-Contents False Belief (Smarties) task than on the Location-Change False Belief (Sally-Anne) task, while the other groups performed similarly on both tasks.

In summary, from comparing the ToM tests, it seems that ToM develops in the same order in ASD as in typical development, and that many children with ASD do show early ToM abilities such as pretend play and understanding perception-knowledge tasks. Next, development will be explored more directly.

To analyse the rate of development and onset of ToM, five separate simple regressions were done (typical development, low-functioning autism, high-functioning autism, Asperger's syndrome, PDD-NOS) with ToM as the dependent variable and Age as the independent variable. The regression results are shown in Table 6.

Table 6.

The Correlation Between Age and Theory of Mind for Typical Development and the Various Autism Spectrum Groups

	Regression	R	R^2	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	95% Confidence Interval for B Lower Bound	95% Confidence Interval for B Upper Bound
TD	Constant	.738***	.545	82.30	7.84				66.24	98.36
	Age			0.61	0.10	.74	5.80	< .0001	0.39	0.82
LFA	Constant	.321	.056	7.42	9.55				-12.57	27.40
	Age			0.17	0.11	.32	1.48	.156	-0.07	0.41
HFA	Constant	.613**	.376	7.73	11.35				-15.82	31.27
	Age			0.61	0.17	.61	3.64	.001	0.26	0.95
AS	Constant	.512*	.263	83.88	17.65				46.95	120.8
	Age			0.51	0.20	.51	2.60	.018	0.10	0.92
PDD-NOS	Constant	.592**	.350	25.13	19.03				-14.85	65.12
	Age			0.79	0.25	.59	3.11	.006	0.26	1.32

Note. Age is given in months; therefore the B coefficient for Age gives the increase in ToM score associated with 1 month's increase in age. 'Constant' gives the average estimated ToM score at 48 months. LFA = low-functioning autism, HFA = high-functioning autism, AS = Asperger's syndrome, PDD-NOS = pervasive developmental disorder not otherwise specified, TD = typical development

Dependent variable: ToM Total Score

**p* < .05, ***p* < .01, ****p* < .001

ToM significantly increased with age in all the groups except for the low-functioning autism group. Overall, a chi-squared test showed no significant difference between the groups in the size of the correlation between age and ToM, $\chi^2 = 4.34$, $p = .362$ (Arsham, n.d.).

Comparison of slopes and intercepts. Fischer's tests were used to compare the groups' slopes, or developmental rate of ToM, and intercepts, or time of onset of ToM abilities. I used 4 years (48 months) as the intercept, as this was the youngest age included in the study. Figure 2 indicates this intercept with a vertical black line.

Overall, there was no significant difference between the slopes of the groups, F (4, 106) = 1.85, p = .125, r = .13. However, from Figure 2 it seemed that the low-functioning autism group in particular deviated from the typically developing group in their rate of ToM development. These two groups were therefore compared, and a significant difference between their slopes, and hence between their rate of development, was found, F (1, 47) = 7.07, p = .011, r = .36.

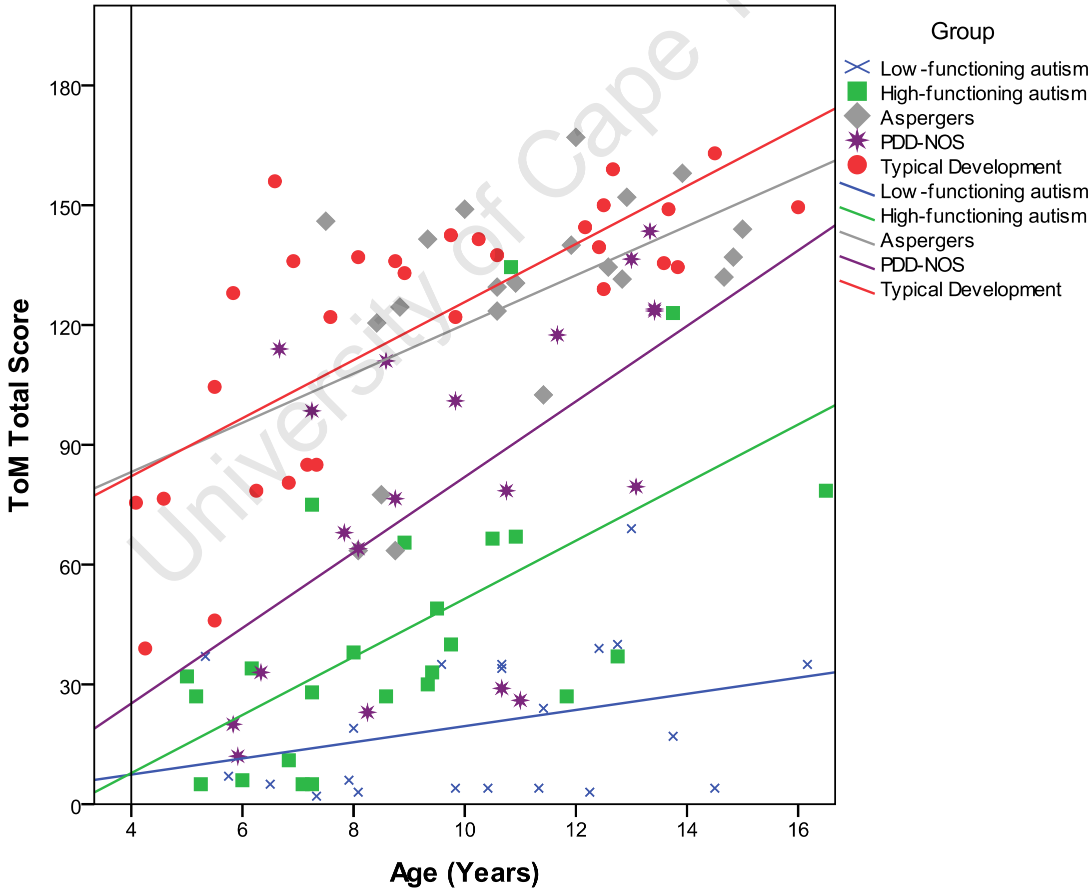

Figure 2. Developmental trajectory of theory of mind. Significant increases in theory of mind were seen in high-functioning autism, Asperger's syndrome, pervasive developmental disorder not otherwise specified and typical development. Overall, there was a similar rate of development between the groups, but different times of onset. The vertical line at 4 years indicates the intercept.

Next, the groups' intercepts were compared. There was a significant difference between the groups' intercepts, or onset of ToM, $F = (4, 110) = 69.64$, $p < .0001$, $r = .62$. As can be seen from Table 6, at 4 years of age the typically developing group had a mean ToM score of 82.30 (*SD* = 7.84), which corresponds roughly to a good understanding of first-order false belief and scattered performance on second-order false belief and understanding non-literal language. The low-functioning autism group had a mean ToM score of 7.42 (*SD* = 9.55) at 4 years, and the high-functioning autism group had a mean score of 7.73 (*SD* = 11.35), both of which indicate an only marginally higher than chance performance on the early ToM battery (understanding desire, understanding seeing-leads-to-knowing, and showing pretend play). The Asperger's syndrome and PDD-NOS groups had mean ToM scores of 83.88 (*SD* = 17.65) and 25.13 (*SD* = 19.03) respectively. This corresponds to an understanding of ToM similar to that of typically developing children in the Asperger's group (F [1, 47] = 0.21, $p = .65$, $r = .07$), and an understanding of the early ToM concepts but not first-order false belief in the PDD-NOS group. However, the regression for the PDD-NOS had a large standard error relative to the other groups, and when looking at the graph it can be observed that very few individual cases in this group perform as predicted by the regression equation. Rather, many cases are similar to that predicted for either Asperger's syndrome or the autism groups.

In summary, the typically developing group and all the ASD groups except for low-functioning autism displayed statistically significant increases in ToM with age. Overall, the groups did not differ significantly in their rate of ToM development; however, the low-functioning autism group showed significantly slower ToM development than the typically developing group. All the ASD groups except for the Asperger's syndrome group showed delayed ToM skills at age 48 months, or 4 years old. From the mean scores at that age, it seems that ToM was especially delayed in the high-functioning and low-functioning autism groups.

Longitudinal study. Ten high-functioning autistic children were followed up one year after the initial assessment to test whether any development of theory of mind took place during this period. A *t*-test showed that the children had significantly improved in ToM from Time 1 to Time 2, t (9) = -2.88, $p = .018$, Hedges' $g = -0.84$. Table 7 shows the pass rate at Time 1 and Time 2. At Time 2, five children passed the Desire task, whereas before none of the 10 had passed it, and three children passed the Perception-Knowledge who had not done so at Time 1. Only 1 child who passed the Perception-Knowledge task at Time 1 failed the task at Time 2. All the children passed the Pretend Play task at both Time 1 and Time 2.

Table 7.

Number of Children Who Passed Each Theory of Mind Task at Time 1 and Time 2

	Pass Year 1		Pass Year 2	
	N	%	*N*	%
Early Module				
Pretend Play	10	100	10	100
Perception-Knowledge	5	50	7	70
Desire	0	0	5	50
Basic Module				
Sticker Hiding	5	50	8	80
Explanation of Action	6	60	5	50
Location-Change False Belief	0	0	4	40
Unexpected-Contents False Belief	0	0	0	0
Intermediate Module				
2nd Order False Belief	-	-	0	0
Strange Stories	-	-	0	0
Advanced Module				
Lies and Jokes	-	-	-	-
Faux Pas	-	-	-	-

Note. The age range for the children who participated in the follow-up study was 5-10 years, with the average age being 8.08 years (at Time 1). Passing criteria were set so that a pass indicated above-chance performance on the task. For specific passing criteria, see the Methods section.

On the basic module, one child passed the Explanation of Action task at Time 2 who had not passed previously; however, 2 children who had passed at Time 1 failed this task at Time 2. At Time 1, none of the children had passed either the Location-Change False Belief or the Unexpected-Contents False Belief tasks. Four children passed the Location-Change task at Time 2. None of the children passed the intermediate or advanced modules.

In contrast to the results of the ToM tests, the children did not do significantly better on the control questions at Time 2, t (9) = -2.25, p = .051, Hedges' g = -0.68. Doing an analysis of covariance with the change in control scores as a covariate, there was still a significant increase in ToM scores from Time 1 to Time 2, M_{T1} = 27.6, M_{T2} = 43.5, F (1,8) = 5.85, p = .042, *partial* η^2 = .423.

In summary, both the longitudinal study and the regression analysis showed severely delayed ToM skills in ASD compared to typical development (with the exception of the Asperger's syndrome group, who performed as well as the typically developing group). The results from the longitudinal investigation also support the results of the regression analysis in showing that ToM development can take place in ASD, and show that increases in ToM can occur in a relatively short space of time. No evidence of a plateau in ToM development was seen within the age range tested. Next, I will explore the different levels of ToM ability found within ASD more fully.

Differences in ToM Ability between ASD and Typical Development

ToM ability within the ASD groups was explored to examine the validity of the current DSM diagnostic categories and to see whether patterns of ToM ability emerged. The groups' scores for the early, basic, intermediate and advanced ToM modules are given in Table 8. As shown in the development section, the typically developing and Asperger's syndrome groups obtained the highest ToM scores, followed by the PDD-NOS group and then the low and high-functioning autism groups. This pattern was true for both the ToM and the control questions, though the groups seemed to perform somewhat better on the control questions than on the ToM questions. The PDD-NOS group had a large variation in ToM scores; much more so than any of the other groups had. From the scores it would seem that the Asperger's syndrome group performed better than the typically developing group. However, it should be kept in mind that the youngest participant in the typically developing group was 4 years old, and thus not expected to pass the intermediate and advanced ToM tasks. The youngest participant in the Asperger's syndrome group was 7 years old, at which age, in typical development, it is expected that the child should be able to pass the intermediate ToM tasks. The excellent performance of the Asperger's syndrome group is thus partly explained by the restricted age range in this group.

Table 8.

Theory of Mind Scores by Difficulty Level

	LFA (*n* = 21)		HFA (*n* = 24)		AS (*n* = 21)		PDD-NOS (*n* = 20)		TD (*n* = 30)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Early Control[a]	5.19	2.02	6.25	2.27	7.95	0.05	7.20	1.51	7.97	0.18
Early ToM[a]	3.43	1.83	5.04	2.03	7.52	1.40	6.05	2.21	7.50	1.36
Basic Control[b]	11.19	12.21	21.24	10.67	32.43	1.81	27.90	4.97	31.80	2.91
Basic ToM[b]	6.67	8.06	11.13	10.58	29.81	3.64	19.35	10.12	30.73	5.30
Intermediate Control[c]	1.62	7.42	8.71	14.35	39.33	5.34	25.00	17.36	33.33	10.68
Intermediate ToM[c]	0.76	3.49	5.92	9.97	32.14	9.08	16.90	12.89	27.85	10.94
Advanced Control[d]	-	-	4.25	14.40	43.62	18.49	21.80	24.88	37.33	27.23
Advanced ToM[d]	-	-	3.00	10.18	33.33	15.11	15.45	18.03	28.63	18.61

Note. In order to compare ToM and control scores, the Pretend and Sticker Hiding tasks, which did not have control questions, were left out of the Early and Basic modules, respectively.
LFA = low-functioning autism, HFA = high-functioning autism, AS = Asperger's syndrome, PDD-NOS = pervasive developmental disorder not otherwise specified, TD = typical development
[a] Maximum score = 8; [b] Maximum score = 33; [c] Maximum score = 33; [d] Maximum score = 56

A one-way analysis of variance showed a significant difference in ToM ability between the groups, F (4,111) = 47.70, $p < .0001$, *partial* $\eta^2 = .632$. *Post hoc* analysis with Bonferroni correction applied revealed significant differences between low-functioning autism and typical development ($p < 0.001$), high-functioning autism and typical development ($p < 0.001$), and PDD-NOS and typical development ($p = 0.001$). There was no difference in ToM between the typically developing and Asperger's syndrome groups ($p = 1.00$). Within ASD, there were significant differences (all $p < .01$) in ToM ability between all the groups except between low-functioning autism and high-functioning autism ($p = .194$). Although there was no significant difference between the low and high-functioning autism groups, there was a trend towards children with high-functioning autism performing better on ToM than children with low-functioning autism (power = .788).

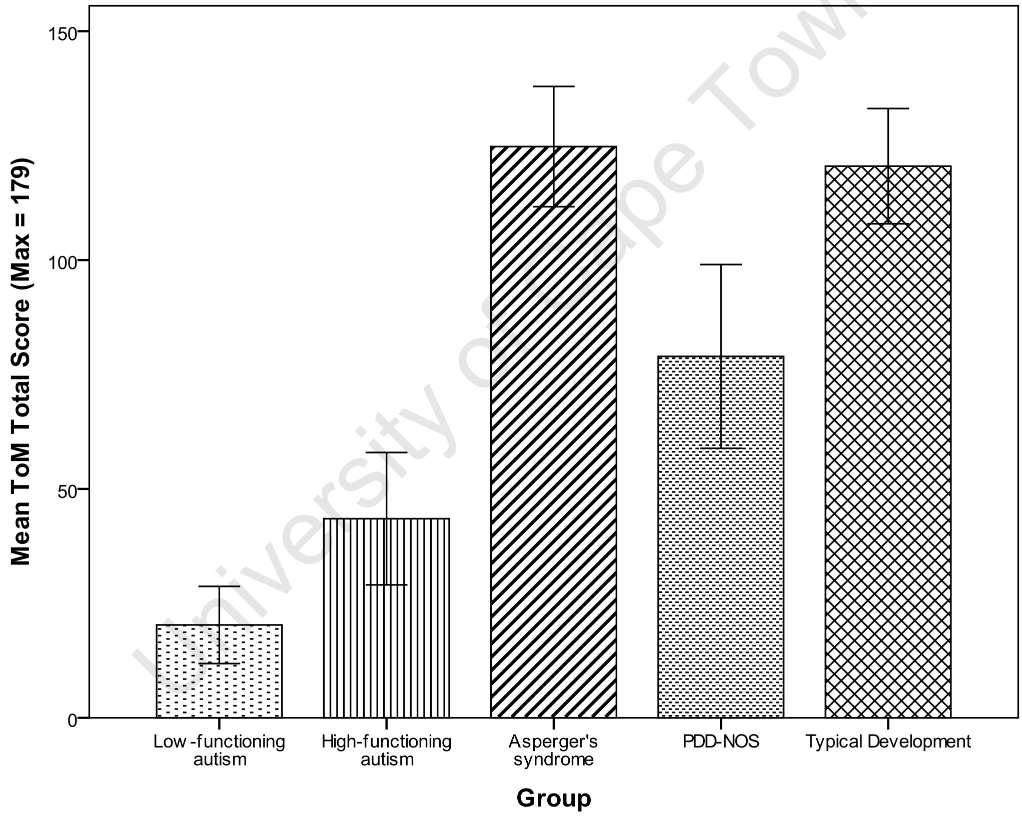

Figure 3. Differences in theory of mind between typical development and the various autism spectrum disorder groups. As shown in the graph, the Asperger's syndrome group performed equivalently to the typically developing control group. ToM performance was increasingly poorer for PDD-NOS, high-functioning autism and low-functioning autism.

Table 9.

Differences Between the Groups in Total Theory of Mind Score

	Group					Significance		
	LFA	HFA	AS	PDD-NOS	TD	*F* (4, 111)	*p*	ESE
n	21	24	21	20	30			
Mean	20.29	43.50	127.05	78.95	120.52	47.70	< .0001	.632
SD	18.52	34.25	28.37	42.89	33.72			

Note. ESE = Effect size estimate *partial* η^2, LFA = low-functioning autism, HFA = high-functioning autism, AS = Asperger's syndrome, PDD-NOS = pervasive developmental disorder not otherwise specified, TD = typical development

Figure 3 and Table 9 reveal large standard deviations in ToM within most of the groups; this shows that there is considerable variation in ToM within the groups. As the large standard deviations found indicate that the diagnostic groups are not homogenous, a cluster analysis was done, using ToM skill as the classifier, to see whether the data clusters into five groups (roughly corresponding to low and high-functioning autism, Asperger's syndrome, PDD-NOS and typical development) or whether a different cluster structure fits the data best.

K-means cluster analysis using Euclidean distances was used to identify groups within the data. Cluster analysis is merely descriptive and cannot prove the existence of groups as different solutions may be obtained with different methods; however, it can be an effective way to classify cases provided that fairly distinct categories naturally occur within the sample (Afifi & Virginia, 1990). First, I examined a five cluster solution, which would be expected if the four ASD subgroups and the typically developing group each had their own ToM profile. The result of the five cluster *K*-means solution is shown in Figure 4 below.

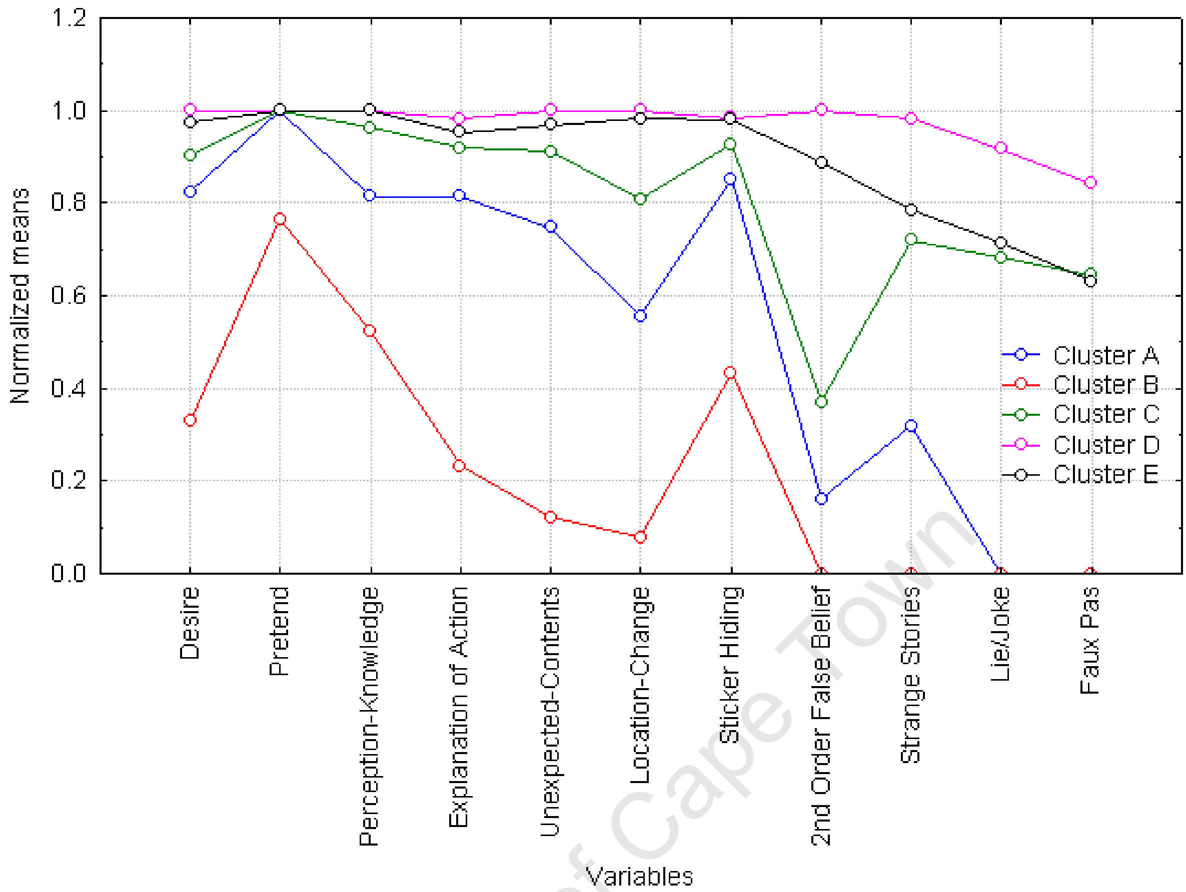

Figure 4. Continuous variable distribution for the five cluster *K*-means analysis. The performance of the cluster on the different ToM tests is shown.

The five cluster solution explained 34% of the variance in total ToM score. Although there were significant differences between the group centroids, this is can be misleading as *K*-means specifically clusters cases so that groups are dissimilar. Clusters A and C and clusters D and E in Figure 4 seem very similar. Therefore, from the continuous variable plot shown in Figure 4 and from inspecting the cluster distributions of the ToM variables, it seems that a three cluster model may be better. A three cluster model also makes sense from the ANOVA results presented above: the Asperger's syndrome and typically developing groups did equally well and the PDD-NOS group did not perform in a uniform way, leaving roughly three groups: low-functioning autism, high-functioning autism/PDD-NOS and Asperger's syndrome/typical development/PDD-NOS. The results of the 3 cluster *K*-means solution is shown in Figure 5. Analysing the cluster distributions and the continuous variable plot, a three cluster solution seems reasonable. Again, there were significant differences between the cluster centroids on all the ToM tasks (all $p < 0.0001$). The three cluster solution explained 44% of the variance in total ToM score. Three clusters also accounted for more of the variance in ToM than two or four clusters (R^2 = .43 and .37 respectively), and proved to be the most interpretable clinically.

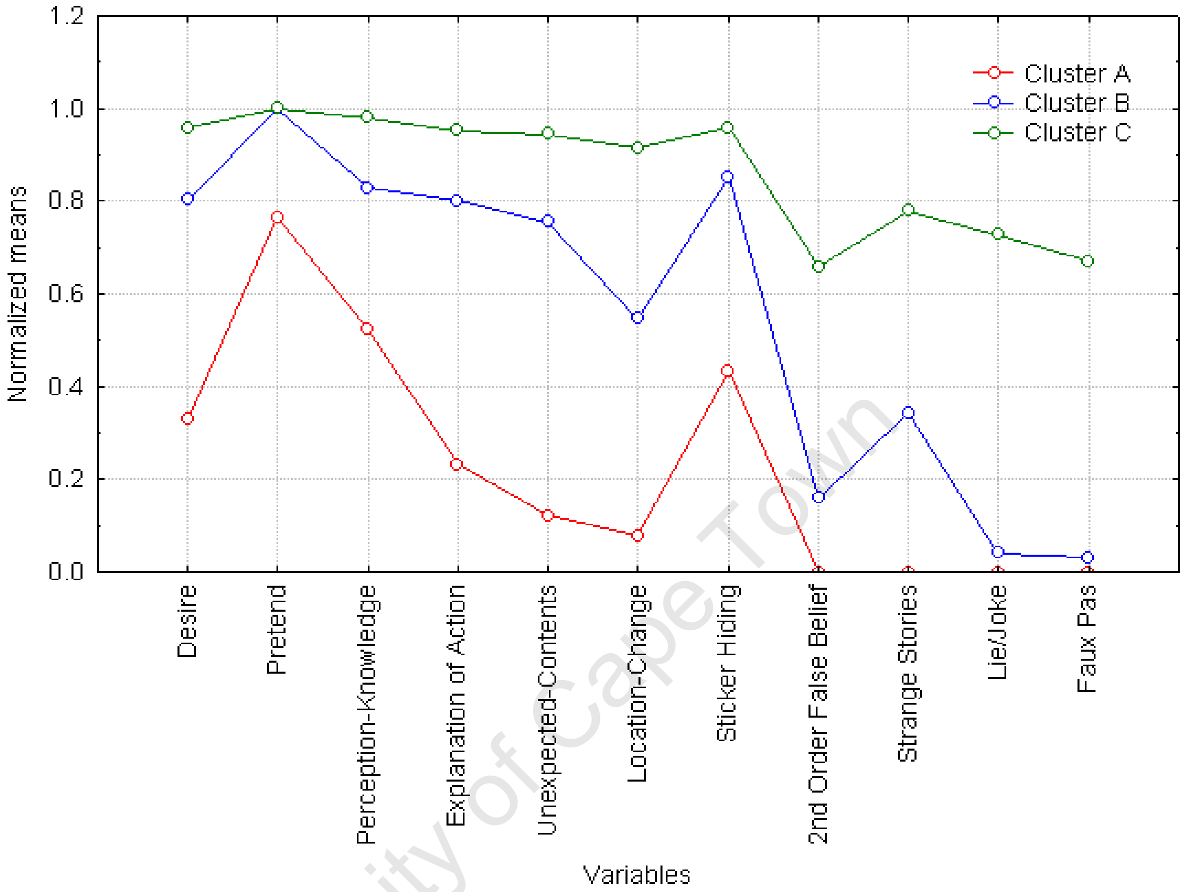

Figure 5. Continuous variable distribution for the three cluster *K*-means analysis. Three distinct clusters emerged. Participants in Cluster A did well on very easy, non-verbal ToM tasks such as Pretend and Sticker Hiding, but poorly on most other ToM tasks. Participants in Cluster B did fairly well on the easy and basic modules, but performed poorly on more difficult ToM tasks. Participants in Cluster C performed well on all ToM tasks.

The three cluster solution produces quite distinct groups. Participants in Cluster A do well on very easy, non-verbal ToM tasks such as Pretend and Sticker Hiding, but poorly on most other ToM tasks. Participants in Cluster B do fairly well on the easy and basic modules, but perform poorly on more difficult ToM tasks. Participants in Cluster C perform well on all ToM tasks. As shown in figures 6 and 7, the Unexpected-Contents and Location-Change false belief tests discriminated best between Cluster A and the other two clusters, and the Lies and Jokes and Faux Pas tests discriminated effectively between Clusters B and C.

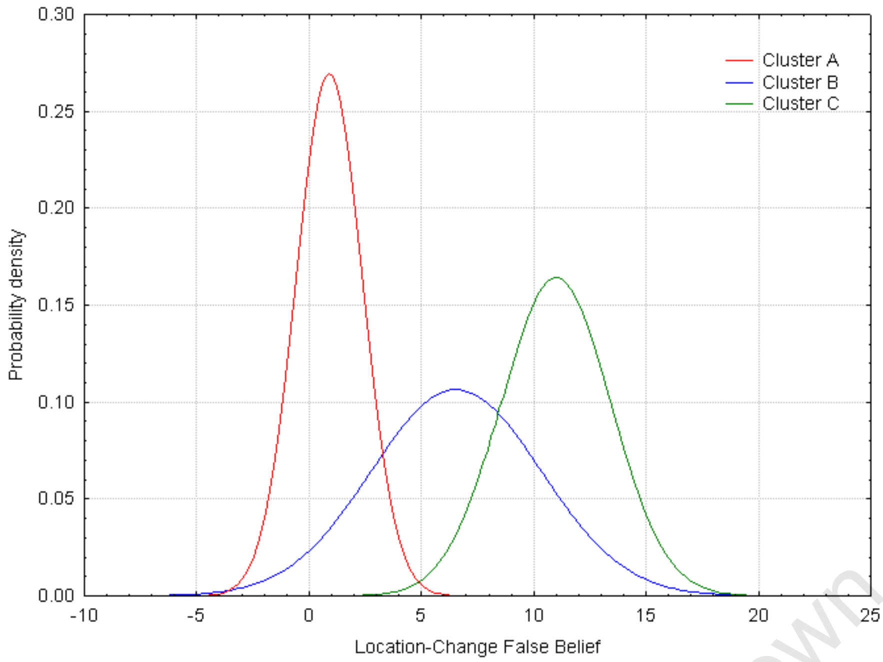

Figure 6. The distribution of performances of the three clusters on the Location-Change False Belief task. Cluster A performed poorly on the Location-Change False Belief task, Cluster C performed very well, and Cluster B performed midway between the other groups.

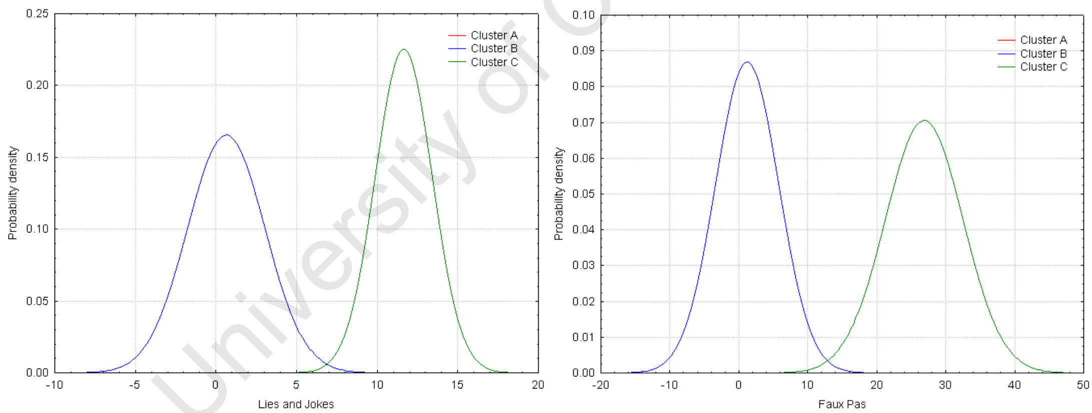

Figure 7. The performance distribution of the clusters on the Lies and Jokes and Faux Pas tasks. Cluster C performed much better than Cluster B on both these tasks. No children from Cluster A proceeded to the Lies and Jokes and Faux Pas tasks.

Table 10 compares the cases in the three clusters on demographic variables. Firstly, it can be observed that the standard deviations for the three clusters are all much smaller than those for the DSM-IV diagnostic groups: The largest standard deviation for the clusters was 25.80, compared with the DSM-IV groups where the largest standard deviation was 42.89, and the smallest was 18.52. Notably, most typically developing children are in Cluster C. Children with Asperger's syndrome either fall within Cluster B or Cluster C, the two clusters associated with greater ToM performance. Children with low and high-functioning autism were grouped in Clusters A and B, with Cluster A containing children with more severe impairments. The typically developing child in Cluster A was one of the youngest children in the sample and very withdrawn, which may explain his grouping with this cluster. The typically developing children in Cluster B were all 7 years of age or younger, were typically non-English first language speakers and of low or medium SES backgrounds. The fact that children with low and high-functioning autism are split between Clusters A and B indicates the often arbitrary nature of the division between low and high-functioning autism. While children on different ends of these groups are very different, children with an IQ score close to 70 (which forms the division between low and high-functioning autism) can be very similar.

Table 10.

The Three Groups Obtained from K-Means Cluster Analysis

	Clusters			Significance	
	A	B	C	F/χ^2	*p*
N	36	29	49		
ToM	23.78 (15.78)	62.84 (25.80)	135.45 (14.27)	688.89	< .0001
FSIQ	70.47 (11.72)	81.76 (14.49)	93.17 (14.69)	28.83	< .0001
Age				11.33	< .0001
Mean (SD)	9.13 (2.94)	8.42 (2.86)	11.22 (2.54)		
Range	4 - 16	4 - 16	5 - 16		
DSM Subtype				64.78	< .0001
LFA	16 (44%)	3 (10%)	-		
HFA	13 (36%)	9 (31%)	2 (4%)		
AS	-	3 (10%)	18 (37%)		
PDD-NOS	6 (17%)	6 (21%)	8 (16%)		
TD	1 (3%)	8 (28%)	21 (43%)		
School Type				48.24	< .0001
Autism specific	24 (67%)	14 (48%)	1 (2%)		
Special Needs / Home schooling	11 (31%)	7 (24%)	21 (43%)		
Mainstream	1 (3%)	8(28%)	27 (55%)		

Note. FSIQ = full scale IQ, SD = standard deviation, LFA = low-functioning autism, HFA = high-functioning autism, AS = Asperger's syndrome, PDD-NOS = pervasive developmental disorder not otherwise specified, TD = typical development

The demographic information presented in Table 10 shows an association between ToM ability and the kind of school environment a child is currently in. Except for the young typically developing children, all children in Clusters A and B were in autism-specific or general special needs schools, with the majority of the children being in an autism-specific school (which mostly caters to children with more severe autism behaviours). In Cluster C, the ASD children were mostly in general special needs schools; in other words, schools that are more academically focused than the autism schools, but where classes are smaller and more structured. The rest of the ASD children in this cluster were in mainstream schools, with the exception of one child who was in an autism-specific school.

In summary, an ANOVA could distinguish between Asperger's syndrome, PDD-NOS and low and high-functioning autism (though the difference between the last two groups was not significant). However, the PDD-NOS group especially, but also the other ASD groups, had high standard deviations on the ToM scores, indicating a lack of homogeneity within the groups. The ANOVA could not distinguish Asperger's syndrome from typical development. A cluster analysis using Euclidean distances between ToM performance split the sample into three clusters corresponding to poor ToM performance (Cluster A), below average ToM performance (Cluster B) and normal ToM task performance (Cluster C).[1]

The effect of cognitive functioning on ToM task performance. Different levels of ability between the groups, for example comprehension ability, general intellectual functioning or executive functioning, may confound group results. The groups were therefore compared on comprehension of ToM tasks and cognitive functioning and ANCOVAs were done to eliminate the effects of confounding variables.

Comprehension of ToM tasks. To test whether the groups understood the instructions and were able to answer questions appropriately, one-way analyses of variance were done to compare the groups' scores on the control questions. Each battery had a similar number of control questions to ToM questions. As many children in the ASD group did not advance to the intermediate and advanced ToM modules, and therefore did not complete all those control questions, the groups' control scores on the separate modules (in which they had to complete every question), rather than their overall control score, were compared. As shown in Table 11, there were significant group differences in control scores on the early, basic and intermediate modules. *Post hoc* analyses with Bonferroni correction applied revealed that the low and

[1] This is not to say that social cognition is not impaired within this group, but rather that the children performed in the normal range on the set of ToM tasks administered.

high-functioning autism groups performed significantly more poorly than the typically developing groups on the early, basic, and intermediate modules' control questions (all $p < .05$). The PDD-NOS group also performed more poorly than the typically developing on the control questions of the basic module.

Table 11.

Differences Between the Groups in Control Scores on the Four Theory of Mind Modules

	Significance			Post hoc comparisons
	F	*p*	ESE	
Early	14.74[a]	< .0001	.347	LFA, HFA < AS, TD
Basic	17.48[b]	< .0001	.419	LFA, HFA, PDD-NOS < AS, TD
Intermediate	3.81[c]	.008	.188	HFA < PDD-NOS, AS, TD*
Advanced	0.47[d]	.708	.027	*ns*

Note. ESE = Effect size estimate *partial* η^2, LFA = low-functioning autism, HFA = high-functioning autism, AS = Asperger's syndrome, PDD-NOS = pervasive developmental disorder not otherwise specified, TD = typical development
* The LFA group is not shown as there was only one participant who completed this module
[a] *df* = 4,111; [b] *df* = 4,97; [c] *df* = 4,66; [d] *df* = 3,50

Because of the PDD-NOS, low and high-functioning autism groups' poor performance on control questions, a series of planned comparisons, shown in Table 12, were done comparing the groups' performances on control and ToM tasks. As the largest group differences were found in the Early and Basic modules, these modules were chosen for comparison to keep the Type I error rate low and a Bonferroni correction was applied to the significance values.

The planned comparisons revealed that these groups still performed significantly lower on the ToM questions than on the control questions (see Table 12). In contrast to this, there were no significant differences between the control and ToM scores for the typically developing group, who performed well on both sets of questions. We would expect this pattern of performance if the ASD groups had specific ToM deficits rather than, or over and above, more general deficits in language comprehension and expression.

Table 12.

Comparison of Control and Theory of Mind Scores on the Early and Basic Modules

Group	Module	Test Type				Significance	
		Control		ToM			
		M	*SD*	*M*	*SD*	*t*	*p*
LFA	Early	5.19	2.02	3.43	1.83	5.10	< .001*
	Basic	11.19	12.21	6.67	8.06	3.01	.003*
HFA	Early	6.25	0.31	5.04	0.36	3.57	< .001*
	Basic	21.13	1.61	11.13	1.67	7.66	< .001*
PDD-NOS	Basic	27.90	4.97	19.35	10.12	5.91	< .001*
TD	Early	7.97	0.27	7.50	0.33	1.57	.1118
	Basic	31.80	1.44	30.73	1.50	0.90	.370

Note. Group performances are only displayed for the modules on which the ASD groups performed significantly more poorly than the typically developing group. The Asperger's syndrome group is not included as this group performed similarly to the typically developing group on all the control tasks.

* Significant at $\alpha = .007$ (Bonferroni correction)

Differences in cognitive functioning. The results of the five one-way analyses of variance for cognitive functioning are shown in Table 13. Significant group differences were found in verbal IQ, verbal generativity, inhibition/set shifting, digit span and processing speed. The PDD-NOS and low and high-functioning autism groups performed more poorly than the typically developing group on verbal IQ, verbal generativity, digit span and processing speed (all $p < .025$). All the groups performed similarly to the typically developing group on the inhibition/set shifting task, except the low-functioning autism group who performed significantly more poorly on this task ($p = .045$). The Asperger's syndrome and the typically developing group performed equally well on all the tasks except for the processing speed task, on which the Asperger's syndrome group performed more poorly than the typically developing group (all $p = .02$). It is of course possible that the single significant result between these two groups could be an artefact, or due to chance. Overall, the typically developing and Asperger's syndrome groups performed the best, the autism groups performed the most poorly and the PDD-NOS group performed midway between the high-functioning autism and Asperger's syndrome groups.

Table 13.

Group Performances on Measures of Cognitive Function

	LFA		HFA		AS		PDD-NOS		TD		Significance Test		
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	*F*	*p*	ESE
Verbal IQ	59.24	5.41	75.42	9.71	98.57	16.54	79.45	12.28	93.47	18.76	28.36[a]	< .001	.505
Verbal Generativity	24.00	10.68	33.67	15.82	48.95	15.38	36.00	12.81	54.29	17.77	10.76[b]	< .001	.341
Inhibition/ Set Shifting	133.00	76.54	118.27	37.76	103.05	32.17	117.73	38.16	86.26	28.82	2.72[c]	.035	.126
Digit Span	2.79	2.46	5.11	2.58	7.86	2.50	5.31	3.26	7.79	2.06	12.12[d]	< .001	.353
Processing Speed	5.33	3.92	8.10	4.51	12.81	4.65	10.00	4.71	17.33	4.97	19.85[e]	< .001	.466

Note. Verbal Generativity represents the number of words generated for semantic and category fluency. The Inhibition/ Set Shifting score is the time taken to complete the Colour-Word Interference task. A high score indicates a poor performance. Digit Span is given as a scaled score with the typical population mean being 10. Processing Speed is given as a scaled score with the typical population mean being 20.
LFA = low-functioning autism, HFA = high-functioning autism, AS = Asperger's syndrome, PDD-NOS = pervasive developmental disorder not otherwise specified, TD = typical development
[a] df = 4,111; [b] df = 4, 83; [c] df = 4, 76; [d] df = 4, 89; [e] df = 4, 91

The cognitive variables measured were all significantly correlated with ToM score ($p < .001$), as shown in Table 14, and thus an analysis of covariance was performed to confirm that differences in ToM score between the groups were not due to any confounding factors. 77 participants were used for this analysis, as 39 participants were either below 6 years and did not do all the tasks, or could not complete some of the tasks. In the latter case, discontinuation was due to two main reasons: inability to read, which excluded these participants from the Colour-Word Interference task, and lack of knowledge of the alphabet or numbers, which excluded them from the Verbal Fluency task.

Table 14.

Correlation Matrix of Theory of Mind and Measures of Cognitive Function

	ToM	Processing Speed	Digit Span	Inhibition/ Set shifting	Verbal Generativit y	VIQ
ToM	--	0.621***	0.577***	-0.388***	0.657***	0.68***
Processing Speed		--	0.494***	-0.313**	0.52***	0.613***
Digit Span			--	-0.235*	0.429***	0.57***
Inhibition/Set shifting				--	-0.303**	-0.239*
Verbal Generativity					--	0.613***
VIQ						--

Note. ToM = theory of mind, VIQ = verbal IQ

* p < .05, ** p < 0.01, *** p < 0.001

The analysis of covariance showed that, even after controlling for cognitive functions associated with ToM performance, there remained a significant difference in ToM between the groups, F (4, 67) = 10.08, $p < .0001$, *partial* $\eta^2 = .376$. The estimated ToM scores, had VIQ, generativity, inhibition, working memory and processing speed been similar amongst the groups, are given in Table 15.

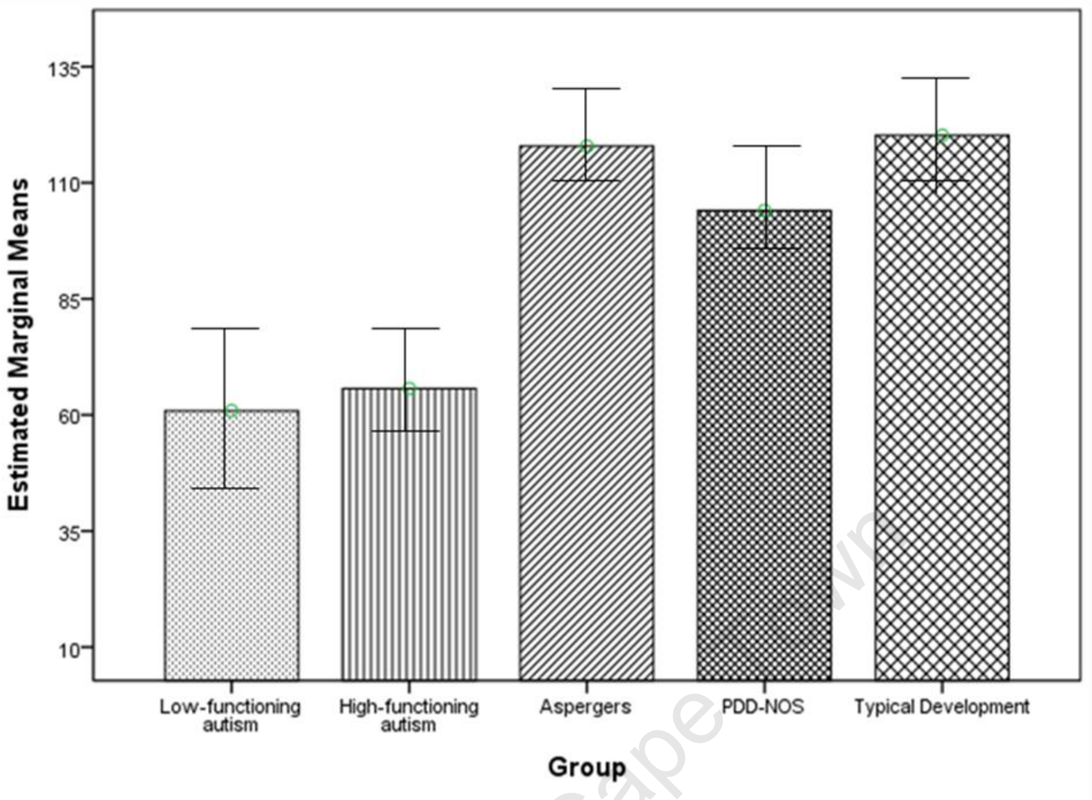

Figure 8. Estimated average ToM scores after controlling for verbal IQ, generativity, inhibition, working memory and processing speed. When the afore-mentioned cognitive factors are statistically controlled for, the low and high-functioning groups perform equally. A significant difference remains between these groups and the Asperger's syndrome and typically developing groups. PDD-NOS performance increases, but remains poorer than the typically developing and Asperger's syndrome groups. *Note:* Covariates appearing in the model are evaluated at the following values: Verbal IQ = 84.62, Verbal Generativity = 43.51, Inhibition/Set Shifting = 108.05, Digit Span = 6.40, Processing Speed = 11.43. Error bars indicate standard error scores.

Post hoc comparisons with Bonferroni correction revealed that significant differences remained between all the ASD groups and the typically developing group (all $p \leq .003$), with the exception of the Asperger's syndrome group. Again, the Asperger's syndrome and the typically developing groups did the best on ToM, followed by the PDD-NOS group, and then the low and high-functioning autism groups.

To summarise, I hypothesized that the Asperger's syndrome and high-functioning autism groups would do equally well on measures of ToM; doing better than the low-functioning autism group, but more poorly than the typically developing group. This hypothesis was only partly upheld: both before and after controlling for confounding cognitive factors such as IQ, the high-functioning autism group did better than the low-functioning autism group but more poorly than the typically developing group. However, a large difference between the performance of the Asperger's syndrome group and the high-functioning group emerged; in fact, the Asperger's syndrome group performed as well on tests of ToM as the typically developing group did. The PDD-NOS group displayed a heterogeneous performance, with average ToM scores between that of the high-functioning autism group and the Asperger's syndrome group. This is illustrated in the results of the cluster analysis, which cluster the sample into a mostly low-functioning group, a mostly high-functioning group, and a group consisting mostly of Asperger's syndrome and typically developing children. Cases from the PDD-NOS group form part of each of the three clusters.

Investigating Diagnostic Criteria: Language Delays and Asperger's Syndrome

To investigate the large differences found between the Asperger's syndrome and high-functioning autism groups, the diagnoses of the participants with Asperger's syndrome were examined. As discussed in the Methods section, the diagnoses of the ASD participants were made by clinicians independent of this study and were used for this study as given. However, when I interviewed parents about these children's early language development, I found that 11 of the 21 children had language delays; defined as no single words at 2 years and phrase speech only after 3 years of age. Therefore, according to the DSM-IV, these children should in fact be diagnosed with high-functioning autism. One of the language delayed participants also had an IQ score in the mildly intellectually disabled range although the DSM-IV specifies that intellectual functioning should be average or above average in Asperger's syndrome (American Psychiatric Association, 2000).

The Speech Delay and No Speech Delay Asperger's syndrome groups were compared on ToM and measures of cognitive function and the results are displayed in Table 16.

Table 15.
Performance of the Speech Delay and No Speech Delay Asperger's Syndrome Groups on Cognitive Measures

	No Speech Delay		Speech Delay		Significance Test		
	Mean	SD	Mean	SD	*t*	*p*	ESE
Age	11.02	2.76	11.59	2.22	-0.54	.600	-0.22
ToM	124.82	32.26	124.77	28.34	0.004	.997	0.00
FSIQ	101.80	12.17	97.18	14.92	0.58	.570	0.32
VIQ	99.09	17.98	95.64	16.87	0.47	.647	0.19
Verbal Generativity	46.82	16.53	50.09	14.34	-0.50	.625	-0.20
Inhibition/ Set Shifting	108.73	41.40	95.27	17.52	0.99	.338	0.41
Digit Span	8.73	2.53	7.09	2.17	1.63	.119	0.67
Processing Speed	14.82	4.09	11.18	4.58	1.96	.064	0.80

Note. 11 Children with Asperger's syndrome formed part of the Speech Delay group and 10 formed part of the No Speech Delay group. ESE = effect size estimate, Hodge's *g*, ToM = theory of mind, FSIQ = full scale IQ, VIQ = verbal IQ

There were no differences between the Speech Delay and No Speech Delay groups in ToM score. Neither were there any differences in age, verbal IQ, inhibition, generativity, working memory or processing speed. Figure 8 illustrates how similar the two groups were on the ToM measures.

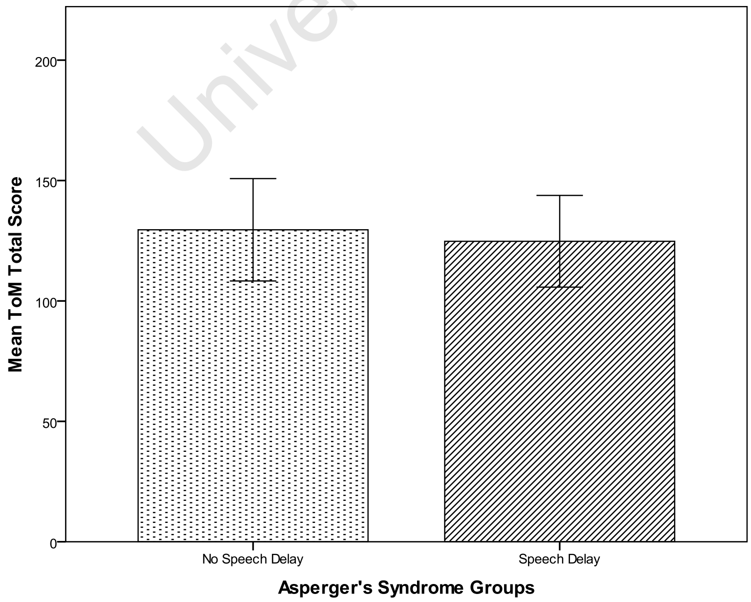

Figure 9. The ToM performance of the Speech Delay and No Speech Delay groups. The error bars denote the 95% confidence intervals.

In light of the disputable diagnosis of 11 participants in the Asperger's syndrome group, the ToM group comparison was repeated, this time with the 11 Speech Delay 'Asperger's Syndrome' participants as part of the high-functioning autism group. Rerunning the analysis of covariance with ToM as the dependent variable, Group as the independent variable (n_{LFA} = 21, n_{HFA} = 35, n_{AS} = 10, $n_{\text{PDD-NOS}}$ = 20, n_{TD} = 30) and verbal IQ, verbal generativity, inhibition/ set shifting, digit span and processing speed as covariates, there was still a significant difference between the groups on ToM, F (4,68) = 3.29, p = .016, *partial* η^2 = .16; however there was no longer a significant difference between high-functioning autism and Asperger's syndrome (p = .550).

In conclusion, distinctive levels of ToM ability did emerge within ASD, but these clusters did not correspond to the current DSM diagnostic criteria. PDD-NOS was particularly problematic in that it did not display a homogenous profile or fit within any specific cluster. The large differences in ToM ability, particularly between autism and Asperger's syndrome, do not support a singular autistic disorder category. Although the differences between high-functioning autism and Asperger's syndrome were no longer significant after children with language delays were removed from the Asperger's group and included in the high-functioning autism group, the cluster analysis still showed large difference in ToM between higher-functioning ASD children. In light of the fact that the Speech Delay Asperger's syndrome group performed equivalently to the No Speech Delay Asperger's syndrome group and very differently from children with high-functioning autism, their inclusion within the Asperger's syndrome group seems valid. Only two children in the high-functioning autism group performed similarly to most of the children with Asperger's syndrome, as can be seen from the cluster analysis results. The difference in school placement and diagnosis between the clusters suggests that ToM can be used to predict adaptive functioning in children with ASD.

Discussion

The aim of the study was to see whether differences in ToM ability and its development would be present in ASD, with the final goal of being able to split the proposed singular DSM-V autistic disorder category into different levels of functioning. With this aim in mind, the development versus stability of ToM in ASD will first be discussed. Differences in ToM between ASD subgroups, both DSM-defined and those identified by cluster analysis, will be discussed next. Lastly, recommendations will be given with regard to diagnostic problems within ASD.

Summary and Implications of Results

ToM development or stability? I investigated whether ToM development is delayed or deviant in ASD. The deviant development hypothesis predicts that ToM never reaches the level seen in typical development, or that ToM develops in a different sequence to that seen in typical development. The specific developmental delay hypothesis predicts that ToM development does occur in ASD, albeit more slowly. Thus the same skills should be seen in ASD as in typical development, though these skills may only be displayed much later in the lifespan than in typically developing individuals.

The performance of the high-functioning autism, PDD-NOS and Asperger's syndrome groups are consistent with the delayed development hypothesis. All the ASD groups except for the Asperger's syndrome group showed delayed onset of ToM skills at 4 years old, with most children not passing age-appropriate false belief tasks at this age. However, comparison of the regression analyses revealed significantly better ToM skills with increasing age in the high-functioning autism, Asperger's syndrome and PDD-NOS groups. There was furthermore no difference in the *rate* of ToM development between these groups and the typically developing group. Neither does the task performance of these ASD groups suggest a deviant *pattern* of ToM skill acquisition. There were no large differences between higher-functioning ASD (high-functioning autism, Asperger's syndrome and PDD-NOS) and typical development in the order in which tasks were passed. There was also no evidence of a plateau in the development of ToM skills. In summary, the higher functioning ASD groups showed a delayed onset in ToM skills, but a normal rate and sequence of development, as predicted by the delayed development hypothesis.

However, there was a disparity between the performance of the higher-functioning ASD groups and that of the low-functioning autism group. ToM in low-functioning autism seems better explained by the deviant development account. In this group, ToM performance did not increase with increasing age. It therefore seems that ToM might develop in higher functioning ASD, but not in ASD with concurrent intellectual disability.

The improvement in ToM skills seen in the 10 children who were reassessed after 1 year also supports the hypothesized delayed development in high-functioning autism. Although these children were extremely deficient in ToM skills (with most children not passing first-order false belief tasks, despite the average age being 8 years) some increases in ToM were observed at the follow-up assessment. Specifically, on the early ToM module, five children passed the Desire task at Time 2, whereas before none of the 10 had passed it, and three children passed the Perception-Knowledge task who had not done so previously. On the basic module, four children passed the Location-Change False Belief task at Time 2, whereas none had passed it at Time 1. With the exception of the performances of one child on the Perception-Knowledge task and two on the Explanation of Action task, if a child passed a task at Time 1, he or she also passed the task at Time 2. This suggests that those who passed genuinely understood the tasks, and therefore that genuine improvement in ToM was observed.

Regarding the inconsistent performance of some children on the ToM tasks, the fact that one boy performed worse on the Perception-Knowledge task at follow-up may be because this task consisted of yes/no answers. It is possible that he merely passed by chance during the first assessment. It is unclear why two children failed the Explanation of Action task who had passed it before. It may be that these children can sometimes use mental state reasoning effectively, but cannot do so consistently or when not concentrating on the task at hand. Another possible explanation is a loss of ToM skills. However, as these three children did similarly or better on other ToM tasks at Time 2, there is no reason to suspect loss of skills in this sample. Therefore, both the regression analysis and the follow-up study suggest development of ToM skills with age in higher-functioning ASD.

In summary, the research results support the delayed development hypothesis for certain ASD subgroups; specifically, high-functioning autism, PDD-NOS and Asperger's syndrome. These groups displayed a normal rate and sequence of ToM development, though the onset of ToM abilities was delayed. Similar to the results of the current study, Steele et al. (2003) and Paynter and Peterson (2009) found that ToM did increase in ASD, and Sparrevohn and Howie (1995) concurred that a normal sequence of ToM development is present in ASD.

In contrast to these results, the longitudinal studies by Ozonoff and McEvoy (1994) and Holroyd and Baron-Cohen (1993) have failed to find ToM development in autism. In the case of Holroyd and Baron-Cohen, ToM was reassessed after 7 years, a much longer time span than the 1 year follow-up presented here. The conflicting results between studies might be explained by examining the ASD population represented in a study. For example, Holroyd and Baron-Cohen used a sample of predominantly low-functioning children. If you compare the performance of the low-functioning group from the current study with that of Holroyd and Baron-Cohen's sample, the results agree: ToM performance did not increase with increasing age. It therefore seems that ToM might develop in higher functioning ASD, but not in ASD with concurrent intellectual disability, which explains the disparate research results. Another explanation for the conflicting results between studies is that the current study, like that of Steele et al. (2003), used younger children and measured earlier developing ToM skills than those assessed by Holroyd and Baron-Cohen and Ozonoff and McEvoy, which made it possible to identify pre-false belief ToM development in certain ASD groups.

In conclusion, a distinction was seen between ToM performance in low-functioning autism and higher-functioning ASD. *Delayed* development of ToM was observed in high-functioning autism, PDD-NOS and Asperger's syndrome. In contrast, ToM did not seem to improve in low-functioning autism, supporting the *deviance* model of development in this group. The differences observed between ASD groups may explain previous contradictory research results.

Age of development of false belief reasoning in ASD. As false belief reasoning has traditionally been the principal test of ToM, its onset within the current ASD sample was investigated. In his paper on specific ToM delays in autism, Baron-Cohen (1989) reported that the youngest age at which children in his sample passed first-order false belief tasks was 11 years. Happé (1995) suggested that individuals with autism require at least a verbal mental age of 9 years 2 months in order to pass false belief tasks. Though mental age was not calculated for this sample, these chronological and mental age estimates are similar to the youngest chronological age that was needed for children with autism and PDD-NOS to pass false belief tasks in this sample: While the majority of typically developing 4-6 years olds and all typically developing children 7 years and older passed first-order false belief tasks, the average age at which children with autism passed both of the first-order false belief tasks was 13.5 years. The youngest child with autism to pass both first-order false belief tasks was 10 years old; the youngest child with PDD-NOS to pass was 8 years old. Of all the ASD children who passed, only two had IQs in the mild intellectual disability range (notably, one of these boys was diagnosed as having Asperger's syndrome rather than autism). These results therefore agree with Happé's suggestion that a certain level of cognitive functioning or language ability is necessary to pass false belief tasks.

Asperger's syndrome: Intrinsic or extrinsic ToM? The Asperger's syndrome group was the only ASD group to show no impairments in ToM: Although there was a difference in ToM onset between the high-functioning autism and PDD-NOS groups and typical development, there was no difference in onset of ToM between Asperger's syndrome and typical development. These results need to be interpreted with caution, for two reasons. Firstly, the youngest participant with Asperger's syndrome was 7 years old. The reason for the narrower age range in the Asperger's syndrome group is that many children with Asperger's syndrome are only diagnosed at around 6 or 7 years because they show less severe behavioural and communication problems than children with autism (Rosenberg et al., 2009). This meant that all the Asperger's syndrome participants in this study could pass at least first-order false belief tasks, so that the onset of ToM abilities, as tested by the battery, had to be extrapolated for this group.

Secondly, although there were no significant differences in ToM found between Asperger's syndrome and typical development, this does not mean that children with Asperger's syndrome have normal social skills. These children's parents reported problems with making friends, turn taking in conversation, understanding non-literal speech, and other problems related to deficits in ToM. From these reports it is unclear whether greater ToM test results are due to improved intrinsic, or true, ToM skills, or due to improved compensatory skills.

The fact that the low-functioning autism group showed no increase in ToM with age, and many higher-functioning ASD children showed excellent ToM task performance while still struggling with real-life social situations, suggests that compensatory skills play a role in the observed increases in ToM. Unlike the other ASD groups who participated in the study, the low-functioning autism group does not have the cognitive capacity to understand what is required from the tests, and therefore may not be able to compensate for their diminished intrinsic ToM by reasoning out the correct answer. Frith et al. (1994) demonstrated that there may be an ASD subgroup who use compensatory skills to pass ToM tasks, but struggle with real life social situations, and an ASD group in which ToM skills truly are more developed. Thus, although ToM impairments in Asperger's syndrome are certainly of a lesser nature than those seen in autism, these children may not be as socially competent as their ToM results would suggest. This also applies to older ASD children whose ToM scores may be higher than they were in early childhood. Clearly, the ToM tests currently used do not measure all aspects of ToM, and are not good simulations of natural social scenes. These tests may be sufficient to detect severe ToM deficits, but are not sensitive enough for children with mild ToM impairments, as is likely to be the case in Asperger's syndrome.

In conclusion, the results support a delayed development hypothesis for high-functioning autism, PDD-NOS and Asperger's syndrome. ToM does develop in these ASD groups, but the onset of ToM is much later than in typical development. The low-functioning group was the only group to perform as predicted by the deviant development hypothesis; the other groups showed a normal rate and sequence of development, with no plateaus in ToM development. The fact that the low-functioning autism group showed no increases in ToM suggests that certain language and cognitive skills may be necessary for ToM development, or that compensatory skills are the cause of the observed increases in ToM performance in the other ASD groups. Though compensatory skills may help on pencil-and-paper tasks and basic interactions, they may not be enough for effective social understanding, as was demonstrated by parents' reports of relatively poor social skills in children with Asperger's syndrome (with high ToM scores) in comparison with typically developing children. Thus, though ToM may improve with age, this does not mean that older children with ASD no longer have social and communicative impairments. This improvement in ToM is also not universal in ASD, as is evident from the performance of the low-functioning autism group.

Differences in ToM ability between the ASD subgroups. One of the study's aims was to establish whether ASD children could be differentiated into distinct groups of ToM ability independent of their general cognitive and executive functioning ability. If no distinct clusters of ToM ability emerged, the results would support the proposed singular autistic disorder category. If clusters similar to the DSM-IV diagnostic categories emerged, this would argue against removing these subgroups from the DSM. If clusters of ToM appeared that did not correspond to the DSM-IV diagnostic categories, this would indicate that a new classification system might be helpful.

When I compared the DSM-IV diagnostic categories, I found large differences in ToM performance between the ASD groups, even after differences in verbal IQ, verbal generativity, inhibition/set shifting, working memory and processing speed were statistically factored out. On average, children with high-functioning autism passed early ToM tasks such as pretend play and understanding desire, but failed to understand first-order false belief tasks. Children with low-functioning autism did not, on average, show pretend play or understand others' desires. Children diagnosed with PDD-NOS performed significantly better on the ToM tasks than children diagnosed with autism, but worse than children with Asperger’s syndrome, who performed similarly to typically developing children. It is interesting that children with PDD-NOS did better on the ToM tasks than children with high-functioning autism despite the fact that these groups’ verbal IQ scores (which are highly correlated with ToM performance) were equivalent. This finding agrees with Baron-Cohen’s (1989) viewpoint that even though higher verbal IQ scores are correlated with higher ToM scores, a high verbal IQ is not sufficient to pass ToM tasks.

Results from previous studies have shown that children with PDD-NOS may have better (Begeer et al., 2003) or worse (Rieffe, Meerum Terwogt, & Stockmann, 2000) ToM skills than children with autism; these divergent results highlight the heterogeneity found in PDD-NOS. In this study too, some children with PDD-NOS performed similarly to children with low or high-functioning autism on ToM tasks, while others showed a much better performance, similar to that seen in Asperger’s syndrome. Given that a goal of diagnosis is to describe a similar set of conditions, these results call into question the use of having a category such as PDD-NOS. This idea will be explored further in the discussion on the results of the cluster analysis.

Contrary to many previous research results, large differences emerged between the high-functioning autism and Asperger’s syndrome groups, even after controlling for cognitive factors that could influence participants' performance on the ToM tasks. This finding concurs with that of four previous studies comparing high-functioning autism and Asperger's syndrome (Kuroda et al., 2011; Ozonoff et al., 1991; Paynter & Peterson, 2009; Ziatas et al., 1998). The reason that Klin (2000) and Spek et al. (2009) did not find differences between these groups may be that their participants were much older than those in the current sample. It is thus possible that their high-functioning autism participants may have “caught up” with their Asperger’s syndrome participants. It could also be argued that these studies used such difficult tests that both the Asperger’s syndrome and high-functioning autism groups did badly on the tasks. Certainly, in the case of Klin’s (2000) study, the Asperger’s syndrome group provided more elaborate answers than the high-functioning autism group, but still performed more poorly than the typically developing group because of irrelevant responses. However, this argument would not hold for Spek et al. (2009) and Dahlgren and Trillingsgaard (1996), who mostly used similar tests to those used in this study, namely, first and second-order false belief, Strange Stories and Faux Pas. A more likely explanation is that studies have found different results because different diagnostic criteria were used to define

Asperger’s syndrome and high-functioning autism. For example, Ozonoff and colleagues (1991) classified children with previous language delays as part of the Asperger’s syndrome group if they did not have any current language impairments. Indeed, in the current study it was found that 11 of the children diagnosed with Asperger's syndrome in fact had a history of language delays (though their language was not currently impaired). When these children were placed in the high-functioning autism group, the distinction between high functioning autism and Asperger's syndrome disappeared. It therefore seems that the reason for the contradictory results is not different methodologies, but different criteria for diagnosis.

Reviewing the results thus far in terms of evaluating both the current and proposed DSM diagnostic criteria, several problems are evident. Firstly, the PDD-NOS group did not have a uniform profile for ToM or general cognitive abilities. Secondly, this study shows that Asperger’s syndrome is not always diagnosed strictly according to DSM-IV criteria and depending on how the diagnosis of Asperger’s syndrome was made, research results may differ considerably. In this study there were clearly two different groups; the vast majority of children in the higher ToM group were diagnosed with Asperger’s syndrome. However, half of those diagnosed as Asperger’s syndrome had early language delays. In strict accordance with the DSM-IV, these children should not have received a diagnosis of Asperger’s syndrome. Hence, a different study may classify these children as high-functioning autism, resulting in better ToM performance in this group, and obscuring the two phenotype clusters found in my study. Thus, studies may get different results depending on how their participants are diagnosed.

Considering that the language delayed and non-delayed Asperger’s syndrome groups were strikingly similar (these groups performed equally well on both ToM and general cognitive functioning tasks) it is my opinion that early language development is not a good criterion on which to base diagnosis. Other researchers have come to similar conclusions regarding early language history. Prior et al. (1998) found different clusters of ASD on the basis of ToM skills and verbal ability, with cluster A being more autism-like and clusters B and C more like Asperger's syndrome and PDD-NOS. Critically, early language development was not associated with any specific cluster. Similarly, Howlin (2003) and Eisenmajer et al. (1998) also found that language delays do not predict symptomatology or social outcome in older individuals with higher-functioning ASD. If early language delay cannot predict differences in cognitive or behavioural outcomes, then it may not be a valid or reliable criterion for separating groups.

With these problems with the diagnostic criteria in mind, a cluster analysis was done to examine ToM differences within ASD by means of grouping similar cases together rather than comparing existing diagnostic categories. Previous cluster analyses have used intellectual functioning, type and severity of symptoms, and social behaviour to identify subgroups within ASD (see Willemsen-Swinkels & Buitelaar, 2002, for a summary). However, to my knowledge this was the first study to use ToM to cluster cases. I suggest it might be better to cluster ASD groups on what is suspected to be a primary cause of the symptoms and social behaviours in ASD, rather than clustering individuals on those more distal factors.

Previous cluster analyses have found from two to four ASD subclassifications, ranging from mild/atypical to severe ASD (Castelloe & Dawson, 1993; Prior et al., 1998; Sevin et al., 1995; Siegel, Anders, Ciaranello, Bienenstock, & Kraemer, 1986; Stevens et al., 2000). The cluster analysis in this study produced three well-differentiated ToM groups. These groups were roughly associated with current DSM diagnosis but not directly related to it. For instance, participants in Cluster A did well on very easy, non-verbal ToM tasks such as Pretend Play and Sticker Hiding, but poorly on most other ToM tasks. These participants were mostly diagnosed with autism, both low and high-functioning. No Asperger's syndrome participants fell into this group and the PDD-NOS participants who were part of this group were either very young (under 6 years old) or had severe autistic symptoms. Cluster B was mostly made up of high-functioning autism and PDD-NOS participants. The rest of Cluster B consisted of the highest functioning children of the low-functioning autism group and the lowest-functioning children with Asperger's syndrome, as well as the youngest typically developing children in the sample. Participants in this cluster did fairly well on the easy and basic ToM modules, but performed poorly on more difficult ToM tasks. Participants in Cluster C performed well on all ToM tasks. This cluster consisted mostly of typically developing children and children with Asperger's syndrome, as well as a few children with PDD-NOS. Only two children with autism (both high-functioning) formed part of this cluster.

The cluster analysis provides a strong argument for classifying children on a dimensional rating scale based on ToM performance, for several reasons. Firstly, the cluster analysis produced groups that were much more homogenous on ToM ability than were the groups based on DSM-IV classification. Secondly, better ToM task performance seems to be strongly correlated with decreased severity of symptoms (as has also been found in Joseph & Tager-Flusberg, 2004; and Lerner et al., 2010). Though this study did not aim to explore the correlation between ToM and symptom severity, a negative relationship is suggested by examining the frequency of the diagnostic categories within the clusters. Children with autism, who have the most severe symptomatology of the ASDs, were mostly placed in the most severely impaired cluster. This cluster was associated with the lowest IQ scores, but did not only contain children with intellectual disability. The cluster analysis also placed the Asperger's syndrome children, as well as children with language delays but with currently strong cognitive capabilities, both generally and in terms of ToM, in the highest ToM ability cluster. Furthermore, it refined the PDD-NOS diagnosis by splitting these children into the three clusters based on their degree of impairment.

Thirdly, and related to symptom severity, the clusters were strongly associated with adaptive functioning, which can be inferred from the type of school environment in which the child is able to be placed. Children in Cluster C, the highest ToM cluster, were mostly in mainstream or general special needs schools. The general special needs schools are characterised by smaller classes than mainstream, but larger classes than autism-specific schools. These schools also have a more academic focus than autism-specific schools. The majority of the children in Clusters A and B were in autism-specific schools, and none of the ASD children in these clusters were in mainstream schools. One limitation is that, because of the sampling method used, this sample was biased towards children in autism-specific schools. However, the findings suggest that classifying children according to ToM ability could be of use to the educational system in deciding to which environment a child would be best suited.

In summary, differences in ToM within ASD were examined in two ways: (1) by comparing DSM-IV diagnostic groups and (2) by comparing groups with similar ToM skills, as identified by a cluster analysis. Comparing DSM-IV subgroups revealed that the Asperger's syndrome group performed similarly to typically developing children in the study, the autism groups displayed the most impaired ToM skills, and the PDD-NOS performed midway between these groups. The heterogeneity of the PDD-NOS group and irregular diagnoses of the high-functioning autism/Asperger's syndrome participants support the DSM committee's decision to abandon the current diagnostic categories.

The cluster analysis identified clusters somewhat different to the DSM-IV diagnostic groups. Three groups were identified: poor ToM performance (Cluster A), which consisted mostly of typical autism, below average ToM performance (Cluster B), which consisted mostly of high-functioning autism and PDD-NOS, and normal ToM task performance (Cluster C), which consisted mostly of Asperger's syndrome and typical development. Thus, despite supporting the DSM decision to change the current diagnostic criteria, the results caution against using a single autistic disorder category, as large differences in ToM were found on both an ANOVA and a cluster analysis. It is therefore of practical importance to separate the groups, as these differences should reflect differences in the core ASD aspects of social and communicative functioning, which in turn affect outcome.

Recommendations

Investigating social competence. As discussed previously, current ToM tests may not be good simulations of social situations and may therefore be too easy for children with relatively well-developed ToM skills. One area of ToM research that critically needs attention is the development of more ecologically valid tasks to assess ToM, or more generally, empathy and social skills. There are two ways to go about this: develop increasingly complex tasks that mimic social situations, such as the 'empathic accuracy' tests devised by Ickes (Ponnet, Buysse, Roeyers, & De Corte, 2005; Ponnet, Roeyers, Buysse, De Clercq, & Van Der Heyden, 2004; Roeyers, Buysse, Ponnet, & Pichal, 2001) or test social inference when the participant's attention is divided between tasks. The second option is to assess more basic impairments in empathy, such as eye contact and social reciprocity, and to assess responses to social situations at a physiological rather than a cognitive level. If autism is indeed at its core a social impairment, all children with ASD should do poorly on such tasks, regardless of their performance on current 'advanced' ToM tests. Another benefit of the latter approach is that this does not exclude children who are non-verbal or intellectually disabled. Another area of study should be behaviours in autism that have no counterpart in normal development, such as obliviousness to parents or caretakers (VanMeter, Fein, Morris, Waterhouse, & Allen, 1997). Many of these behaviours have not been studied as comprehensively as traditional ToM skills, but play a critical role in the empathic deficits seen in autism.

Diagnosis. Several problems with the current diagnostic criteria emerged. Firstly, my results emphasize that the category PDD-NOS is too heterogeneous to be of much use. Currently, it merely serves as a wastebasket for any atypical cases and is therefore not useful as an indicator of the severity of symptoms, what type of intervention may be needed, or of prognosis. PDD-NOS as a "lesser variant of autism" (Matson & Boisjoli, 2007; Serra et al., 2002) may be true for some of the cases that fall in this category, but individual cases may present with severe intellectual disability or symptomatology. Secondly, this study found no differences in ToM or general cognitive functioning based on early language development. Possibly because of this, I found that clinicians do not diagnose Asperger's syndrome strictly according to the criterion of normal language development. Clearly, if diagnoses are made subjectively rather than according to recognised international standards, this can cause some confusion among researchers, clinicians and the general public.

From these results, different courses of action are immediately evident: (1) change the diagnostic criteria for Asperger's syndrome so that they do not require the absence of language delays and create different categories within PDD-NOS, or (2) abandon the diagnostic categories of Asperger's syndrome and PDD-NOS and create a singular autistic disorder category.

Regarding to the first strategy, PDD-NOS may be split into several clusters within which presentation is similar, as has been proposed by Walker et al. (2004). Asperger's syndrome could be diagnosed according to present day language and ToM skills, rather than early language development. This change in diagnostic criteria seems feasible given the lack of evidence for a correlation between previous language delays and current functioning. Indeed, though recent opinion holds that a diagnosis of Asperger's syndrome is impossible if language development is delayed (American Psychiatric Association, 2000) this was not always the prevailing opinion. Many researchers have held the opinion that Asperger's syndrome can be diagnosed in cases with a history of language delay as long as current language ability is at an age-appropriate level (Cederlund & Gillberg, 2007; Tantam, 1988; Wing, 1981), and that children who displayed autism-like symptoms in early childhood could develop language and social skills to the extent that they appeared more Asperger-like. With an increasing number of children throughout the world who have access to early, intensive intervention strategies, this may be occurring more and more frequently.

A problem with maintaining separate diagnostic categories within ASD is that we do not know whether autism, Asperger's syndrome and PDD-NOS have separate aetiologies. However, as Szatmari and colleagues argue, "if the distinctiveness of disorders depends on demonstrating separate aetiologies, then there are few disorders in child psychology that could be truly separate" (Szatmari, Archer, Fisman, Streiner, & Wilson, 1995, p. 1670). Practically, the ability to identify groups with different levels of cognitive and behavioural functioning is clinically more useful. Such differences between high-functioning-autism, PDD-NOS and Asperger's syndrome were shown in the current study.

A larger obstacle to retaining separate diagnoses for Asperger's syndrome, autism and PDD-NOS, is access to service provision and resources. If diagnosis is linked to symptom severity, this causes difficulty for clinical trials and access to medical aid. For example, the US' Food and Drug Administration (FDA) will not approve a drug's use for PDD-NOS if the drug is approved for autism, even though the same children may have the two different diagnoses at different ages (Swedo et al., 2008). In certain states in the US, medical aid will pay for therapy for children with autism, but not for children with PDD-NOS (Lord, 2010). In light of such problems, one diagnosis for all children showing the triad of social and communicative deficits and repetitive behaviours would be favourable.

The DSM-V planning committee has proposed that autism, Asperger's syndrome and PDD-NOS should be grouped into a single category in the DSM-V. However, the results of this study show that large differences in ToM exist within ASD. I would like to suggest that a singular ASD category would suffer from the same problems that are currently associated with PDD-NOS; namely that the range of symptom severities is too big for the diagnostic category to be of any practical use. This would in all likelihood mean that teachers, clinicians and researchers would revert to the old diagnostic names to explain a child's 'autism phenotype', with the resulting classification problems that go along with those diagnoses.

To counter such problems, it has been proposed that a dimensional classification system be created within DSM-V autistic disorder. The rating system currently proposed was shown in Table 1 (p. 15) and rates severity of autistic symptoms within two areas: social-communication skills and presence of fixed interests and repetitive behaviours. This approach was taken as several researchers are now of the opinion that ASD is multidimensional and that the core areas in ASD (socialization, communication and repetitive and restricted behaviours) may be separately affected and thus unevenly impaired (for instance, diagnosis may be complicated if a great deal of restricted and repetitive behaviours are present, but social-communicative skills are very good; Happé & Ronald, 2008; Swedo et al., 2008). Lord, Leventhal and Cook (2001) have suggested that ASD diagnosis should include even more dimensions. These authors suggested that classification should include nonverbal intellectual ability, expressive language, and social-emotional reciprocity with repetitive behaviours/restricted interests. However, assessment of severity may become unduly complicated when many domains have to be rated. As has been previously discussed, the proposed dimensional categories for the DSM-V are currently only vaguely defined and are therefore vulnerable to misinterpretation or irregular application.

However, a multidimensional classification system may not be necessary. There is some suggestion that intellectual functioning, social cognition and repetitive behaviours covary. Children with high IQs tend to be more socially active and have obsessive interests and rituals, whereas severely intellectually disabled children tend to be the most socially withdrawn and show more repetitive behaviours and motor stereotypies (Fombonne et al., 1994). Thus, perhaps only a minority of cases may feature significantly different levels of severity in the core impairments.

The results of the cluster analysis indicated how a dimensional classification could be derived using ToM as classifier. This analysis revealed that classification according to ToM ability seems to be related to symptom severity and adaptive functioning (as has previously been reported by Fombonne et al., 1994; Frith et al., 1994; Lerner et al., 2010), making it useful as an indicator of level-of-functioning. The cluster analysis also suggests which tests would be most useful in order to best distinguish the categories on a shorter ToM test battery. The Unexpected-Contents and Location-Change False Belief tests discriminated best between Cluster A and the other two clusters, while the Lies and Jokes and Faux Pas tests discriminated effectively between Clusters B and C.

To conclude, the problems associated with the current diagnostic criteria have made it necessary to reconsider the current ASD categories. Firstly, the heterogeneity found within PDD-NOS reduces the usefulness of having such a category; secondly, the current DSM criteria do not allow for developmental growth; in other words, the possibility that some children first diagnosed with autism may develop into a more Asperger-like picture. A better classification system is therefore needed to distinguish between differences within ASD. A single autistic disorder category would emphasise the similarity between the ASDs and potentially create better service delivery, but (like PDD-NOS) may be too broad to be practically useful. Regarding the large ToM differences observed between high-functioning autism and Asperger's syndrome, I suggest that some distinction between different levels of ability needs to be made if these diagnostic categories are to be grouped into one in the DSM-V, especially since social cognition is a core feature of ASD. Two options, namely (1) changing the diagnostic criteria for Asperger's syndrome and PDD-NOS, and (2) creating levels-of-functioning within autistic disorder were discussed. Regarding option 2, problems remain with a dimensional rating system based on a single factor; however, a multidimensional system may be difficult to implement in practise. A single factor classification based on ToM may be useful to identify children at different levels of functioning. In identifying the severity of symptoms and adaptive functioning of an individual, it is hoped that such a classification would be able to indicate best-practise intervention, predict outcome, and chart social-communicative development.

General Comments on the ToM Battery. The age ranges at which different ToM skills were present in the typically developing group were roughly consistent with previously reported results from the UK and USA (Brüne & Brüne-Cohrs, 2006; Sullivan et al., 1994; Wellman & Woolley, 1990). However, some ToM tasks were found to be more difficult than others that were in the same module. For instance, all the groups found the Desire task much more difficult than the other tasks in the early module, and more difficult even than some of the basic ToM tasks such as Sticker Hiding and Explanation of Action. Steele et al. (2003) used the same set of tests in their early and basic ToM modules, and found that the Desire task, although more difficult than the Pretend Play task, was easier than the Perception-Knowledge and False Belief tasks from the basic module. These contrasting results may mean that cultural bias or some difference in the presentation of the task might have made the Desire task difficult for this South African sample.

Many of the children also found the Second-Order False Belief task more difficult than Strange Stories, though previous research suggested that second-order false belief is necessary to understand the difference between many non-literal statements, such as the difference between a lie and an ironic statement (Ackerman, 1981; M. R. Pollio & H. R. Pollio, 1979). This seemingly paradoxical result may be due to the fact that the Strange Stories task contains a number of different false statements – lies, jokes, pretending, double bluff, persuasion and irony. The double bluff story was the only one that required participants to explicitly state the second-order belief processes ("the prisoner *knows* that the other army *thinks* that he will lie, so he tells the truth to prevent them finding the tanks"). The other statements did not require such explicit explanations. The pretend play statement in this task merely required children to understand the distinction between reality and imagination (which typically develops at around 3-4 years) and was easily understood by even the youngest typically developing children in the sample. Most children also understood *lying* – in fact, in young and autistic children "s/he was lying" was the most common response, even when this response was not appropriate for the question. Thus, children may have found the Second-Order False Belief task more difficult than Strange Stories because of the range of easier stories included within the latter test.

Interestingly, the low and high-functioning autism groups performed much better on the Unexpected-Contents False Belief (Smarties) task than on the Location-Change False Belief (Sally-Anne) task, while the other groups performed similarly on both tasks. The largest difference between autism and the other groups was that the autism groups had much lower verbal IQ scores (though the difference between high-functioning autism and PDD-NOS was negligible). Therefore a possible explanation for the poorer performance on the Location-Change task is that it required giving explanations, whereas the Unexpected-Contents task simply required one-word answers. Another possible explanation for the poor performance on the Location-Change task is its presentation as a picture story, whereas real objects were shown in the Unexpected-Contents task. Perhaps the autism groups found it easier to reason about others' beliefs if the object on which the false belief was based was physically present; in other words, the picture story might have added another level of abstraction that needed to be understood to make inferences about mental states. However, the results of Wellman et al.'s (2001) meta-analysis do not support any of these arguments: they found that neither the type of false belief task presented, nor the medium in which the tasks were presented, nor what type of response was required (pointing, yes/no answers, or longer verbalisations) made any differences in the age at which children passed false belief tasks. However, it should be noted that these authors only included typically developing children in their study, so it is possible that these results cannot be generalised to children with autism. Similar to the results of the meta-analysis, the typically developing children in this study performed equally well on the different false belief tasks.

In summary, the overall performance of typically developing South African children was similar to that found in previous international research. However, performance differed from what was expected for some of the tasks on the battery. These results suggest that the order or content of the battery may need to be revised in future research. More research into the reasons for the autism groups' inconsistent performance on the different false belief tasks would be useful.

Limitations and Directions for Future Research

Home language. There were significant differences in home language between the groups. However, although there were differences in home language, all the children were assessed in their language of schooling (either Afrikaans or English). Therefore, dissimilarities in home language should not have affected test results. An interesting observation was that the PDD-NOS and Asperger's syndrome groups were almost exclusively English speaking. The group differences found may therefore be an indication that many children with ASD are either unidentified or misdiagnosed in certain language groups. Future research is needed to examine the prevalence of ASD in South Africa in diverse cultural groups.

Diagnosis. The diagnosis of ASD was made independently of this study by several different clinicians. It is possible that these clinicians' diagnoses do not correspond to each other – it is well documented that considerable disagreement exists between clinicians, especially concerning the diagnostic categories Asperger's syndrome and PDD-NOS (Howlin & Asgharian, 1999; Williams et al., 2008). This was a limitation that was borne out of necessity, as I am not qualified to make a diagnosis of autism myself, and it was not financially or practically possible to hire a clinical psychologist to individually diagnose all the children who participated in the study. I aim to address this limitation in future research by incorporating diagnostic tools such as the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2001) and the Autism Diagnostic Interview – Revised (ADI-R; Lord, Rutter, & LeCouteur, 2003).

Language and cognitive barriers in the ASD group. The study aimed to include early-developing theory of mind tasks and tests adapted for participants with limited expressive language in order to gain access to the low-functioning autism population. Task adaptations included non-verbal tests, such as Pretend Play and Sticker Hiding, and tests given in multiple choice formats where participants simply had to point to the answer. However, though the tests were adapted to make them accessible for children with autism, the tasks proved to be too cognitively demanding for many low-functioning and young autistic children.

Only 47% of enrolled children with ASD under the age of 7 years were able to complete the tasks. These children are therefore not representative of 'typical' ability or development, but rather represent the most high-functioning children in this age range. Similarly, only 32% of children with low-functioning autism, regardless of their age, were able to complete the tasks. Thus, the low-functioning sample represents a minority within this diagnostic category who have the cognitive ability to complete ToM tasks. The problems most often encountered were that the child's receptive language was too poor to be able to follow instructions, the child was unable to choose an option out of the multiple choice layout, and that the child did not mimic the instructor, and therefore could not complete the Sticker Hiding task.

Additionally, some caution should be taken when interpreting the ToM performance of those children with autism who could participate in the research. There were significant between-group differences in the number of control questions answered correctly. The low and high-functioning autism groups, who only managed to get half of the control questions right, did particularly poorly. Although these groups performed poorly on the control questions, their performance on ToM questions was significantly poorer. In contrast to this, the typically developing group did equally well on the control and ToM questions. These results indicate that although the ASD groups' performances may have been influenced by their comprehension skills, their performance on theory of mind tasks cannot be attributed solely to language ability. For example, ToM deficits were also observed on the relatively non-verbal Sticker Hiding task. Many children, especially those with low-functioning autism, did not understand the concept of 'hiding' at all, and would simply take the sticker for themselves when it was handed to them with the instruction to hide it. This behaviour illustrated the general lack of social reciprocity found within this group.

Regarding the cluster analysis, a fourth group, consisting of those children who did not have the cognitive skills to be able to participate in this study, could speculatively be added to the three clusters identified. Although this group could not participate in the research, they are important to keep in consideration as long-term prognosis is the least favourable for children with severe intellectual disability and absence of language (Nordin & Gillberg, 1998). To be able to include more ASD children in research, future ToM projects need to focus on the development of simple, non-verbal tasks, which require only limited cognitive skills. For example, useful skills to assess include greeting behaviour, eye contact, response to calls and responsiveness to social cues from the examiner. These tasks assess basic or precursor ToM abilities such as joint attention, and do not require extensive language or cognitive skills.

Culturally inappropriate test materials. Some of the content of, and the language used in, the more advanced ToM tasks was not culturally relevant to a South African sample. For example, in one of the Strange Stories scenarios, a character threatens to drown kittens that she cannot keep. The story makes it clear that the character does not wish to hurt the kittens and it is therefore assumed that she is lying in an attempt to persuade the protagonist to buy a kitten. However, in the South African context, especially in rural and township areas, unwanted animals are considered pests and the animal rights movement is less prominent than in the UK. Hence, many children believed that the story character was telling the truth. This can be seen as a different cultural viewpoint on treatment of unwanted animals, rather than a failing of ToM per se. Another example was the Lies and Jokes task, where the protagonist told their parent or teacher something that is not true. Participants had to say whether the protagonist (1) lied to prevent getting caught, or (2) made a joke because he or she was embarrassed. Though the reasoning behind the statement was alluded to in the question (i.e., the child is embarrassed because the parent knows he or she did something wrong or the child will lie to prevent the parent from knowing that he/ she did something wrong), very few children ever said that the child was joking. It is possible that participants in this sample tended to interpret the child's false statement as a lie rather than as a joke because a joke seemed inappropriate in a setting where the child is talking to an authority figure.

Although obviously foreign elements (such as baseball bats) and language were changed beforehand to make the stories more accessible for a South African population, an in-depth adaptation needs to be done, specifically with the advanced ToM battery, and some stories may need to be removed altogether and replaced with new stories.

Follow-up study. A limitation of the follow-up study was that an alternate battery was not used from Time 1 to Time 2, so that it is possible that the children could have learnt the correct answers and that this falsely increased their scores at Time 2. However, none of the examiners ever gave the children the correct answers to the tasks. The strongest argument against the possibility that seeing the tests twice increased children's performance at Time 2 is that children did not perform better on the control questions at Time 2, even though their answers on these questions were corrected at Time 1.

Other limitations of the study are the short follow-up period and small number of children who were reassessed. The significant change in ToM that was observed over the period of 1-year suggests that development can take place and that it can occur over a relatively short space of time. Future research should focus on longitudinal studies of ToM in both low and high-functioning ASD. Furthermore, longitudinal studies with large sample sizes would be beneficial.

Summary and Conclusions

This study investigated the universality and stability of ToM within the subtypes of ASD. Understanding how ToM varies across the lifespan and between ASD subtypes is important for understanding aetiology, for planning effective intervention strategies and for diagnostic purposes. An overarching aim of this study was to see whether specific subgroups of ASD, based on ToM ability, emerged to validate the diagnostic changes proposed for the DSM-V. Understanding ToM within ASD is particularly pertinent as it gives an indication of social competence, one of the core deficits in this disorder.

The main result of this study was that differences in ToM ability emerged between the different subgroups of ASD, both cross-sectionally and in the development of ToM. Low-functioning autism displayed the most impaired ToM skills, followed by high-functioning autism and then PDD-NOS. The Asperger's syndrome group performed equivalently to typically developing children in the study. Although the low-functioning autism group did not display any increase in ToM with increasing age, the high-functioning autism, PDD-NOS and Asperger's syndrome groups all displayed greater ToM skills with age. At least in higher-functioning ASD, these results are consistent with the delayed development hypothesis: it was found that onset of ToM skills in these groups is severely delayed (except in Asperger's syndrome), but that ToM seems to develop at a normal rate. In comparison, the lack of ToM improvement in low-functioning autism is in agreement with the deviant development hypothesis. These findings reveal that a single explanation for ToM deficits may not suffice for the entire autism spectrum.

In terms of the DSM, the results highlight two major problems with the DSM-IV diagnostic criteria. Firstly, the PDD-NOS group did not have a uniform profile for ToM or general cognitive abilities. Secondly, this study shows that Asperger's syndrome is not always diagnosed strictly to DSM-IV criteria. Depending on the way in which the diagnosis of Asperger's syndrome was made, research results – in this case ToM performance – may differ considerably. As has been shown in this study, large differences in cognitive functioning can be found within ASD; however, the presence or absence of early language delay makes little difference in current functioning. The debate surrounding cognitive and behavioural similarities or dissimilarities between high-functioning autism and Asperger's syndrome has lasted for 20 years and no conclusions have yet been reached. We will remain at an impasse until a solution has been reached regarding diagnosis.

Seeing that differences in general cognitive functioning, ToM and executive functioning can make a significant impact on an individual's ability to complete activities of daily living, I recommend that the criteria for Asperger's syndrome and PDD-NOS be revised. If the ASD subgroups are to be grouped into a single autistic disorder category, dimensional categories within this diagnosis need to be made. A cluster analysis showed that ToM may be a useful way in which to create such dimensional categories. Three distinct ToM clusters were found: poor ToM, consisting mostly of children with autism; below average ToM, consisting of children with high-functioning autism and PDD-NOS; and high ToM, consisting mostly of typically developing children and children with Asperger's syndrome.

Dimensional categorization would ideally reflect the level of impairment in the core areas of ASD and be an indication of adaptive functioning, prognosis and eventually pathology. ToM, which I use in the broad sense of the word to include early developing ToM and basic social reciprocity skills, may be a useful way in which to establish such levels-of-functioning as it reflects the core deficits in social competence found in ASD. This study has shown that ToM ability can discriminate different groups within ASD. By replicating this research on larger samples, it will also be possible to establish development norms for the different level-of-functioning subtypes, as has been done in typical development.

Differentiating children by level-of-functioning can contribute greatly to intervention strategies. The results of this study suggest that different intervention strategies for children with low-functioning autism, high-functioning autism and Asperger's syndrome may be more useful than a single social skills training approach for all the ASDs. The low-functioning group may benefit most from imaginary play opportunities and training in recognizing basic emotions. For the high-functioning group it may be most beneficial to learn certain ToM skills in a rote manner, for example understanding that other people's thoughts and feelings may differ from their own and learning some figures of speech and their meaning. Children with Asperger's syndrome may have reasonably well developed skills to cognitively reason about what another person may mean when they say a certain phrase, but may need intervention on how to apply this knowledge to a social situation, such as talking with friends, meeting new people and learning socially acceptable conduct.

Similarly, the finding that ToM has the potential to improve affects the way we think about autism across the lifespan and should be kept in mind when designing interventions for ASD individuals of different ages. However, this does not mean that these individuals do not have lifelong social communication difficulties. There has recently been a shift from question-and-answer type ToM tasks towards examining eye contact, joint attention, and emotional reciprocity. By studying such basic affective tasks, we may be able to identify primitive empathic deficits that I argue to be present in ASD, even in individuals with relatively strong ToM skills. Such a research avenue may also explain why there is an 'autism picture' when ToM deficits are not specific to ASD. To conclude, the areas of social reciprocity and empathy hold much promise for broadening our understanding of ToM and of ASD.

References

Ackerman, B. P. (1981). Young children's understanding of a speaker's intentional use of a false utterance. *Developmental Psychology, 17*(4), 472-480. doi:10.1037/0012-1649.17.4.472

Afifi, A. A., & Virginia, C. (1990). *Computer-aided multivariate analysis* (2nd ed.). New York, NY: Van Nostrand Reinhold.

American Psychiatric Assocation. (n.d.). Childhood Disintegrative Disorder. *DSM-5 Development*. Retrieved January 12, 2011, from http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=96#

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

American Psychiatric Association. (n.d.). Rett's Disorder. *DSM-5 Development*. Retrieved January 12, 2011, from http://www.dsm5.org/ProposedRevisions/Pages/proposedrevision.aspx?rid=95#

Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., & Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature Genetics, 23*(2), 185-188. doi:10.1038/13810

Arsham, H. (n.d.). Test for equality of several correlation coefficients. Retrieved December 14, 2010, from http://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/MultiCorr.htm#rmenu

Astington, J. W. (2001). The future of theory-of-mind research: Understanding motivational states, the role of language, and real-world consequences. *Child Development, 72*(3), 685-687. doi:10.1111/1467-8624.00305

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35*(5), 1311-

1320. doi:10.1037//0012-1649.35.5.1311

Avis, J., & Harris, P. L. (1991). Belief-desire reasoning among Baka children: Evidence for a universal conception of mind. *Child Development*, *62*(3), 460-467. doi:10.2307/1131123

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, *368*, 210-215. doi:10.1016/S0140-6736(06)69041-7

Baron-Cohen, S. (1989). The autistic child's theory of mind: a case of specific developmental delay. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *30*(2), 285-297. doi: 10.1016/0010-0277(85)90022-8

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*, 37-46. doi:10.1016/0010-0277(85)90022-8

Baron-Cohen, S., O'Riordan, M., Stone, V. E., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, *29*, 407-418. doi:10.1023/A:1023035012436

Baron-Cohen, S., Ring, H. A., Wheelright, S., Bullmore, E. T., Brammer, M. J., Simmons, A., & Williams, S. C. R. (2008). Social intelligence in the normal and autistic brain: an fMRI study. *European Journal of Neuroscience*, *11*(6), 1891-1898.

Baron-Cohen, S., & Swettenham, J. (1997). Theory of mind in autism: its relationship to executive function and central coherence. In *D. Cohen & F. Volkmar (Eds.), Handbook of autism and pervasive developmental disorders* (2nd ed., pp. 880-893). New York: Wiley. Retrieved from www.autismresearchcentre.com/docs/papers/1997_BCetal_hypothesis.pdf

Bauminger, N., & Kasari, C. (1999). Theory of mind in high-functioning children with autism. *Journal of Autism and Developmental Disorders*, *29*(1), 81-86. doi:10.1023/A:1025974701090

Begeer, S., Rieffe, C., Meerum Terwogt, M., & Stockmann, L. (2003). Theory of mind–based action in children from the autism spectrum. *Journal of Autism and Developmental Disorders*, *33*, 479-487. doi:10.1023/A:1025875311062

Bibby, H., & McDonald, S. (2005). Theory of mind after traumatic brain injury. *Neuropsychologia*, *43*, 99-114. doi:10.1016/j.neuropsychologia.2004.04.027

Bosacki, S., & Astington, J. W. (2001). Theory of mind in preadolescence: Relations between social understanding and social competence. *Social Development*, *8*(2), 237-255. doi:10.1111/1467-9507.00093

Brüne, M., & Brüne-Cohrs, U. (2006). Theory of mind—evolution, ontogeny, brain mechanisms and psychopathology. *Neuroscience and Biobehavioral Reviews*, *30*, 437–455. Retrieved from www.elsevier.com/locate/neubiorev

Buitelaar, J. K., van der Wees, M., Swaab-Barneveld, H., & van der Gaag, R. (1999). Verbal memory and performance IQ predict theory of mind and emotion recognition ability in children with autistic spectrum disorders and in psychiatric control children. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *40*(6), 869-881. doi:10.1111/1469-7610.00505

Buitelaar, J. K., van der Wees, M., Swaab–Barneveld, H., & van der Gaag, R. (1999). Theory of mind and emotion-recognition functioning in autistic spectrum disorders and in psychiatric control and normal children. *Development and Psychopathology*, *11*(1), 39-58. doi:10.1017/S0954579499001947

Burack, J. A., & Volkmar, F. (1992). Development of low- and high-functioning autistic children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *33*(3), 607-616.

Byrne, R. W., & Whiten, A. (1988). *Machiavellian intelligence: social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford, UK: Oxford University Press.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032-1053. doi:10.1111/1467-8624.00333

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development*, *69*(3), 672-691. doi:10.1111/j.1467-8624.1998.00672.x

Castelloe, P., & Dawson, G. (1993). Subclassification of children with autism and pervasive developmental disorder: a questionnaire based on Wing's subgrouping scheme. *Journal of Autism and Developmental Disorders*, *23*(2), 229-241. doi:10.1007/BF01046217

Cederlund, M., & Gillberg, C. (2007). One hundred males with Asperger syndrome: a clinical study of background and associated factors. *Developmental Medicine and Child Neurology*, *46*(10), 652-660. doi:10.1111/j.1469-8749.2004.tb00977.x

Chakrabarti, S., & Fombonne, E. (2005). Pervasive developmental disorders in preschool children: Confirmation of high prevalence. *American Journal Psychiatry*, *162*(6), 1133-1141. doi:10.1176/appi.ajp.162.6.1133

Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A., & Drew, A. (2000). Testing joint attention, imitation and play as infancy precursors to language and theory of mind. *Cognitive Development, 15*, 481-498. doi:10.1016/S0885-2014(01)00037-5

Dahlgren, S. O., & Trillingsgaard, A. (1996). Theory of mind in non-retarded children with autism and Asperger's syndrome. A research note. *Journal of Child Psychology and*

Psychiatry, *37*(6), 759-763. doi:10.1111/j.1469-7610.1996.tb01469.x

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system manual*. Oxford, UK: Pearson Assessment.

Dennett, D. C. (1989). *The intentional stance*. Cambridge, MA: MIT Press.

Devries, R. (1970). The development of role-taking as reflected by behaviour of bright, average, and retarded children in a social guessing game. *Child Development*, *41*(3), 759-770. doi:10.1111/1467-8624

Dissanayake, C., & Macintosh, K. E. (2003). Mind reading and social functioning in children with autistic disorder and Asperger's disorder. In B. Repacholi & V. Slaughter (Eds.), *Individual differences in theory of mind: implications for typical and atypical development*. Hove, E. Sussex: Psychology Press.

Dyck, M. J., Ferguson, K., & Shochet, I. M. (2001). Do autism spectrum disorders differ from each other and from non-spectrum disorders on emotion recognition tests? *European Child and Adolescent Psychiatry*, *10*(2), 105-116. doi:10.1007/s007870170033

Eisenmajer, R., Prior, M., Leekam, S., Wing, L., Gould, J., Weaver, L. K., & Ong, B. (1996). Comparison of clinical symptoms in autism and Asperger's disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*(11), 1523-1531.

Eisenmajer, R., Prior, M., Leekam, S., Wing, L., Ong, B., Gould, J., & Welham, M. (1998). Delayed language onset as a predictor of clinical symptoms in pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *28*(6), 527-533. doi:10.1023/A:1026004212375

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. doi:10.3758/BRM.41.4.1149

Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition*, *44*, 283-296.

doi:10.1016/0010-0277(92)90004-2

Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: an update. *Journal of Autism and Developmental Disorders*, *33*(4), 365-382. doi:10.1023/A:1025054610557

Fombonne, E., Siddons, F., Achard, S., Frith, U., & Happé, F. (1994). Adaptive behaviour and theory of mind in autism. *European Child and Adolescent Psychiatry*, *3*(3), 176-186. doi:10.1007/BF02720324

Frith, U. (2004). Emanuel Miller lecture: Confusions and controversies about Asperger syndrome. *Journal of Child Psychology and Psychiatry*, *45*(4), 672-686. doi:10.1111/j.1469-7610.2004.00262.x

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transaction of the Royal Society of London. B*, *358*, 459-473. doi:10.1098/rstb.2002.1218

Frith, U., Happé, F., & Siddons, F. (1994). Autism and theory of mind in everyday life. *Social Development*, *3*(2), 108-124. doi:10.1111/j.1467-9507.1994.tb00031.x

Ghaziuddin, M. (2008). Defining the Behavioral Phenotype of Asperger Syndrome. *Journal of Autism and Developmental Disorders*, *38*(1), 138-142. doi:10.1007/s10803-007-0371-7

Greig, T. C., Bryson, G. J., & Bell, M. D. (2004). Theory of mind performance in schizophrenia: diagnostic, symptom, and neuropsychological correlates. *The Journal of Nervous and Mental Disease*, *192*(1), 12-18. doi:10.1097/01.nmd.0000105995.67947.fc

Happé, F. (1993). Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition*, *48*, 101-119. doi:10.1234/12345678

Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders, 24,* 129-154.

Happé, F. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66,* 843-855. Terieved from http://www.jstor.org/stable/1131954

Happé, F., & Frith, U. (1996). The neuropsychology of autism. *Brain, 119,* 1377-1400. doi:10.1093/brain/119.4.1377

Happé, F., & Ronald, A. (2008). The 'fractionable autism triad': A review of evidence from behavioural, genetic, cognitive and neural research. *Neuropsychology Review, 18*(4), 287-304. doi:10.1007/s11065-008-9076-8

Henry, J. D., Phillips, L. H., Crawford, J. R., Ietswaart, M., & Summers, F. (2006). Theory of mind following traumatic brain injury: The role of emotion recognition and executive dysfunction. *Neuropsychologia, 44,* 1623-1628. doi:10.1016/j.neuropsychologia.2006.03.020

Holroyd, S., & Baron-Cohen, S. (1993). How far can people with autism go in developing a theory of mind? *Journal of Autism and Developmental Disorders, 23,* 379-385.

Howlin, P. (2003). Outcome in high-functioning adults with autism with and without early language delays: Implications for the differentiation between autism and Asperger syndrome. *Journal of Autism and Developmental Disorders, 33*(1), 3-13. doi:10.1023/A:1022270118899

Howlin, P., & Asgharian, A. (1999). The diagnosis of autism and Asperger syndrome: findings from a survey of 770 families. *Developmental Medicine & Child Neurology, 41*(12), 834-839. doi:10.1111/j.1469-8749.1999.tb00550.x

Hughes, C. (1998). Executive function in preschoolers: Links with theory of mind and verbal ability. *British Journal of Developmental Psychology, 16*(2), 233-253.

Joseph, R. M., & Tager-Flusberg, H. (2004). The relationship of theory of mind and executive functions to symptom type and severity in children with autism. *Development and Psychopathology*, *16*(1), 137-55. doi:PMC1201455

Kaland, N., Møller-Nielsen, A., Callesen, K., Mortensen, E. L., Gottlieb, D., & Smith, L. (2002). A new 'advanced' test of theory of mind: evidence from children and adolescents with Asperger syndrome. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *43*, 517-528. doi:10.1111/1469-7610.00042

Kaland, N., Smith, L., & Mortensen. (2007). Response times of children and adolescents with Asperger Syndrome on an 'advanced' test of theory of mind. *Journal of Autism and Developmental Disorders*, *37*, 197-209. doi:10.1007/s10803-006-0152-8

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2*, 217-250. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4880460

Kavanaugh, R., Eizenman, D., & Harris, P. (1997). Young children's understanding of pretense expressions of independent agency. *Developmental Psychology*, *33*(5), 764-770. doi:10.1037//0012-1649.33.5.764

Klin, A. (2000). Attributing social meaning to ambiguous visual stimuli in higher-functioning autism and Asperger syndrome: The social attribution task. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, *41*(07), 831-846. doi:10.1111/1469-7610.00671

Kogan, M. D., Blumberg, S. J., Schieve, L. A., Boyle, C. A., Perrin, J. M., Ghandour, R. M., Singh, G. K., et al. (2009). Prevalence of parent-reported diagnosis of autism spectrum disorder among children in the US, 2007. *Pediatrics*, published online Oct 5, 2009. doi:10.1542/peds.2009-1522

Kozinetz, C. A., Skender, M. L., MacNaughton, N., Almes, M. J., Schultz, R. J., Percy, A. K., & Glaze, D. G. (1993). Epidemiology of Rett syndrome: A population-based registry.

Pediatrics, *91*(2), 445-450. Retrieved from http://pediatrics.aappublications.org/cgi/content/abstract/91/2/445

Kuroda, M., Wakabayashi, A., Uchiyama, T., Yoshida, Y., Koyama, T., & Kamio, Y. (2011). Determining differences in social cognition between high-functioning autistic disorder and other pervasive developmental disorders using new advanced "mind-reading" tasks. *Research in Autism Spectrum Disorders*, *5,* 554-561. doi:10.1016/j.rasd.2010.06.022

Lerner, M. D., Hutchins, T. L., & Prelock, P. A. (2010). Brief Report: Preliminary evaluation of the Theory of Mind Inventory and its relationship to measures of social skills. *Journal of Autism and Developmental Disorders*. doi:10.1007/s10803-010-1066-z

Leslie, A. M. (1987). Pretense and representation: The origins of “theory of mind”. *Psychological Review, 94,* 412-426. doi:doi:10.1037//0033-295X.94.4.412

Leslie, A. M., & Frith, U. (1990). Prospects for a cognitive neuropsychology of autism: Hobson's choice. *Psychological Review*, *97*, 122-131.

Lord, C. (2010). Towards DSM V - New trends in diagnosis and classification of ASD. Presented at the IX International Conference Autism-Europe: A Future for Autism, Catania, Italy. Retrieved from http://www.autismeurope2010.org/green_room.html

Lord, C., Leventhal, B. L., & Cook, E. H. (2001). Quantifying the phenotype in autism spectrum disorders. *American Journal of Medical Genetics*, *105*(1), 36-38. doi:10.1002/1096-8628(20010108)105:1<36::AID-AJMG1053>3.3.CO;2-W

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2001). *Autism Diagnostic Observation Schedule (ADOS) manual*. Los Angeles, CA: Western Psychological Services. Retrieved from http://scholar.google.co.za.ezproxy.uct.ac.za/scholar?hl=en&q=Lord%2C+C.%2C+Rutter%2C+M.%2C+DiLavore%2C+P.%2C+%26+Risi%2C+S.+%282001%29.+Autism

+Diagnostic+Observation+Schedule+%28ADOS%29+manual.+Los+Angeles%2C+CA%3A+Western+Psychological+Services&btnG=Search&as_sdt=2000&as_ylo=&as_vis=0

Lord, C., Rutter, M., & LeCouteur, A. (2003). *Autism Diagnostic Interview - Revised*. Los Angeles, CA: Western Psychological Services.

Macintosh, K., & Dissanayake, C. (2006). Social skills and problem behaviours in school aged children with high-functioning autism and Asperger's disorder. *Journal of Autism and Developmental Disorders, 36*(8), 1065-1076. doi:10.1007/s10803-006-0139-5

Manjiviona, J., & Prior, M. (1995). Comparison of Asperger syndrome and high-functioning autistic children on a test of motor impairment. *Journal of Autism and Developmental Disorders, 25*(1), 23-39. doi:10.1007/BF02178165

Matson, J. L., & Boisjoli, J. A. (2007). Differential diagnosis of PDDNOS in children. *Research in Autism Spectrum Disorders, 1*(1), 75-84. doi:10.1016/j.rasd.2006.09.001

Mayes, S. D., Calahoun, S. L., & Crites, D. L. (2000). Does DSM-IV Asperger's disorder exist? *Journal of Abnormal Child Psychology, 29*(3), 263-271. doi:0.1023/A:1010337916636

Meyer, J. A., & Minshew, N. J. (2002). An update on neurocognitive profiles in Asperger syndrome and high-functioning autism. *Focus on Autism and Other Developmental Disabilities, 17*(3), 152 -160. doi:10.1177/10883576020170030501

Miller, J. N., & Ozonoff, S. (2000). The external validity of Asperger disorder: Lack of evidence from the domain of neuropsychology. *Journal of Abnormal Psychology, 109*(2), 227-238. doi:10.1037//0021-843X.109.2.227

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child*

Development, *78*, 622-646. doi:0.1111/j.1467-8624.2007.01018.x

Motulsky, H. (1999). *Analyzing data with graphPad prism* (3rd ed.). GraphPad Software.

Myer, L., Stein, D. J., Grimsrud, A., Seedat, S., & Williams, D. R. (2008). Social determinants of psychological distress in a nationally-representative sample of South African adults. *Social Science & Medicine*, *66*(8), 1828-1840. doi:10.1016/j.socscimed.2008.01.025

Naito, M., Komatsu, S., & Fuke, T. (1994). Normal and autistic children's understanding of their own and others' false belief: a study from Japan. *British Journal of Developmental Psychology*, *12*, 403-416.

Nordin, V., & Gillberg, C. (1998). The long-term course of autistic disorders: update on follow-up studies. *Acta Psychiatrica Scandinavica*, *97*(2), 99-108. doi:10.1111/j.1600-0447.1998.tb09970.x

Noterdaeme, M., Wriedt, E., & Höhne, C. (2009). Asperger’s syndrome and high-functioning autism: language, motor and cognitive profiles. *European Child and Adolescent Psychiatry*, *19*(6), 475-481. doi:10.1007/s00787-009-0057-0

Ozonoff, S., & Jensen, J. (1999). Specific executive function profiles in three neurodevelopmental disorders. *Journal of Autism and Developmental Disorders*, *29*, 171-177. doi:10.1023/A:1023052913110

Ozonoff, S., & McEvoy, R. E. (1994). A longitudinal study of executive function and theory of mind development in autism. *Development and Psychopathology*, *6*, 415-431. doi:10.1017/S0954579400006027

Ozonoff, S., Rogers, S. J., & Pennington, B. F. (1991). Asperger's syndrome: Evidence of an empirical distinction from high-functioning autism. *Journal of Child Psychology and Psychiatry*, *32*(7), 1107-1122. doi:10.1111/j.1469-7610.1991.tb00352.x

Paynter, J., & Peterson, C. (2009). Language and ToM development in autism versus

Asperger syndrome: Contrasting influences of syntactic versus lexical/semantic maturity. *Research in Autism Spectrum Disorders*. doi:10.1016/j.rasd.2009.10.005

Pellicano, E. (2007). Links between theory of mind and executive function in young children with autism: Clues to developmental primacy. *Developmental Psychology*, *43*, 974-990. doi:10.1037/0012-1649.43.4.974

Perner, J., Kain, W., & Barchfeld, P. (2002). Executive control and higher-order theory of mind in children at risk of ADHD. *Infant and Child Development*, *11*(2), 141. doi:10.1002/icd.302

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, *3*(9), 337-344. doi:10.1016/S1364-6613(99)01362-5

Perner, J., Leekam, S., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*, 125-137.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that..." attribution of second-order beliefs by 5- to 10-year-old children. *Journal of Experimental Child Psychology*, *39*(3), 437-471. doi:10.1016/0022-0965(85)90051-7

Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development*, *76*, 502-517. doi:10.1111/j.1467-8624.2005.00859.x

Pollio, M. R., & Pollio, H. R. (1979). A test of metaphoric comprehension and some preliminary developmental data. *Journal of Child Language*, *6*(01), 111-120. doi:10.1017/S0305000900007650

Ponnet, K., Buysse, A., Roeyers, H., & De Corte, K. (2005). Empathic accuracy in adults with a pervasive developmental disorder during an unstructured conversation with a typically developing stranger. *Journal of Autism and Developmental Disorders*, *35*(5), 585-600. doi:10.1007/s10803-005-0003-z

Ponnet, K. S., Roeyers, H., Buysse, A., De Clercq, A., & Van Der Heyden, E. (2004). Advanced mind-reading in adults with Asperger syndrome. *Autism, 8*(3), 249 -266. doi:10.1177/1362361304045214

Pratt, C., & Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development, 61*, 973-982. doi:10.1111/j.1467-8624.1990.tb02835.x

Prior, M., Leekam, S., Ong, B., Eisenmajer, R., Wing, L., Gould, J., & Dowe, D. (1998). Are there subgroups within the autistic spectrum? A cluster analysis of a group of children with autistic spectrum disorders. *Journal of Child Psychology and Psychiatry, 39*(6), 893-902. doi:10.1111/1469-7610.00389

Repacholi, B., & Slaughter, V. (2003). *Individual differences in theory of mind: implications for typical and atypical development*. Hove, E. Sussex: Psychology Press.

Rieffe, C., Meerum Terwogt, M., & Stockmann, L. (2000). Understanding atypical emotions among children with autism. *Journal of Autism and Developmental Disorders, 30*(3), 195-203. doi:10.1023/A:1005540417877

Roeyers, H., Buysse, A., Ponnet, K., & Pichal, B. (2001). Advancing advanced mind-reading tests: Empathic accuracy in adults with a pervasive developmental disorder. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(02), 271-278. doi:10.1017/S0021963001006680

Rosenberg, R. E., Daniels, A. M., Law, J. K., Law, P. A., & Kaufmann, W. E. (2009). Trends in autism spectrum disorder diagnoses: 1994–2007. *Journal of Autism and Developmental Disorders, 39*(8), 1099-1111. doi:10.1007/s10803-009-0723-6

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: links between preschoolers' theory of mind and semantic development. *Child Development*, *72*(4), 1054-1070. doi:10.1111/1467-8624.00334

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind. *Psychological Science*, *17*(1), 74 -81. doi:10.1111/j.1467-9280.2005.01667.x

Saltzman, J., Strauss, E., Hunter, M., & Archibald, S. (2000). Theory of mind and executive functions in normal human aging and Parkinson's disease. *Journal of the International Neuropsychological Society*, *6*(07), 781-788. doi:10.1017/S1355617700677056

Sattler, J. (2008). *Resource Guide to Accompany Assessment of Children : Cognitive Foundations* (5th ed.). San Diego, CA: Author.

Schmitt, A. J., & Wodrich, D. L. (2004). Validation of a Developmental Neuropsychological Assessment (NEPSY) through comparison of neurological, scholastic concerns, and control groups. *Archives of Clinical Neuropsychology*, *19*(8), 1077-1093. doi:10.1016/j.acn.2004.02.002

Schneider, W., Schumann-Hengsteler, R., & Sodian, B. (2005). *Young children's cognitive development: interrelationships among executive functioning, working memory, verbal ability, and theory of mind*. Mahwah, NJ: Routledge.

Serra, M., Jackson, A. E., van Geert, P., & Minderaa, R. B. (1998). Brief report: interpretation of facial expressions, postures and gestures in children with a pervasive developmental disorder not otherwise specified. *Journal of Autism and Developmental Disorders, 28*(3), 257-263. doi:10.1023/A:1026077623805

Serra, M., Loth, F. L., van Geert, P., Hurkens, E., & Minderaa, R. B. (2002). Theory of mind in children with 'lesser variants' of autism: a longitudinal study. *Journal of Child Psychology and Psychiatry*, *43*(7), 885-900. doi:10.1111/1469-7610.00104

Serra, M., Minderaa, R. B., van Geert, P., Jackson, A. E., Althaus, M., & Til, R. (1995). Emotional role-taking abilities of children with a pervasive developmental disorder not otherwise specified. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 36*(3), 475-490. doi:10.1111/j.1469-7610.1995.tb01304.x

Sevin, J. A., Matson, J. L., Coe, D., Love, S. R., Matese, M. J., & Benavidez, D. A. (1995). Empirically derived subtypes of pervasive developmental disorders: a cluster analytic study. *Journal of Autism and Developmental Disorders*, *25*(6), 561-578. doi:10.1007/BF02178188

Sicotte, C., & Stemberger, R. M. (1999). Do children with PDDNOS have a theory of mind? *Journal of Autism & Developmental Disorders, 29*(3), 225-233. doi:10.1023/A:1023032122489

Siegel, B., Anders, T. F., Ciaranello, R. D., Bienenstock, B., & Kraemer, H. C. (1986). Empirically derived subclassification of the autistic syndrome. *Journal of Autism and Developmental Disorders*, *16*(3), 275-293. doi:10.1007/BF01531660

Slade, L., & Ruffman, T. (2005). How language does (and does not) relate to theory of mind: A longitudinal study of syntax, semantics, working memory and false belief. *The British Journal of Developmental Psychology*, *23*(1), 117. doi:10.1348/026151004X21332

Sparrevohn, R., & Howie, P. M. (1995). Theory of mind in children with autistic disorder: evidence of developmental progression and the role of verbal ability. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 36*(2), 249-263. doi:10.1111/j.1469-7610.1995.tb01823.x

Spek, A. A., Scholte, E. M., & Berckelaer-Onnes, I. A. (2009). ToM in adults with HFA and Asperger syndrome. *Journal of Autism and Developmental Disorders*, *40*, 280-289.

SPSS. (2010). *PASW Statistics*. Chicago, IL: Author.

Statistics South Africa. (2006). *Provincial profile 2004: Western Cape* (No. 00-91-01 (2004)). Pretoria, South Africa.

StatSoft. (2010). *STATISTICA version 9*. Tulsa, OK: Author.

Steele, S., Joseph, R. M., & Tager-Flusberg, H. (2003). Developmental change in theory of mind abilities in children with autism. *Journal of Autism and Developmental Disorders, 33*, 461-467. doi:10.1023/A:1025075115100

Stevens, M. C., Fein, D., Dunn, M., Allen, D., Waterhouse, L., Feinstein, C., & Rapin, I. (2000). Subgroups of children with autism by cluster analysis: a longitudinal examination. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*(3), 346-352. doi:10.1097/00004583-200003000-00017

Stone, V. E., Baron-Cohen, S., & Knight, R. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10*, 640-656. doi:10.1162/089892998562942

Stroop, J. R. (1935/1992). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *121*(1), 15-23.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology*, *30*(3), 395-402. doi:10.1037//0012-1649.30.3.395

Swedo, S., Thorsen, P., & Pine, D. (Chairs). (2008). Autism and Other Pervasive Developmental Disorders Conference. Presented at the American Psychiatric Institute for Research and Education [APIRE] conference series "The Future of Psychiatric Diagnosis: Refining the Research Agenda", M.I.N.D Institute, Sacramento, California. Retrieved from http://www.dsm5.org/Research/Pages/AutismandOtherPervasiveDevelopmentalDisordersConference%28February3-5,2008%29.aspx

Szatmari, P., Archer, L., Fisman, S., Streiner, D. L., & Wilson, F. (1995). Asperger's syndrome and autism: Differences in behavior, cognition, and adaptive functioning.

Journal of the American Academy of Child & Adolescent Psychiatry, *34*(12), 1662-1671. doi:10.1097/00004583-199512000-00017

Tager-Flusberg, H. (2007). Evaluating the theory-of-mind hypothesis of autism. *Current Directions in Psychological Science*, *16*, 311-315. doi:10.1111/j.1467-8721.2007.00527.x

Tantam, D. (1988). Asperger's syndrome. *Journal of child psychology and psychiatry and allied disciplines*, *29*(3), 245-255. doi:10.1111/j.1469-7610.1988.tb00713.x

Thede, L., & Coolidge, F. (2007). Psychological and neurobehavioral comparisons of children with Asperger's disorder versus high-functioning autism. *Journal of Autism and Developmental Disorders*, *37*(5), 847-854. doi:10.1007/s10803-006-0212-0

Travis, L., Sigman, M., & Ruskin, E. (2001). Links between social understanding and social behavior in verbally able children with autism. *Journal of Autism and Developmental Disorders*, *31*(2), 119-130. doi: 10.1023/A:1010705912731

VanMeter, L., Fein, D., Morris, R., Waterhouse, L., & Allen, D. (1997). Delay versus deviance in autistic social behavior. *Journal of Autism and Developmental Disorders*, *27*(5), 557-569. doi:10.1023/A:1025830110640

Verté, S., Geurts, H. M., Roeyers, H., Oosterlaan, J., & Sergeant, J. A. (2006). Executive functioning in children with an autism spectrum disorder: Can we differentiate within the spectrum? *Journal of Autism and Developmental Disorders*, *36*(3), 351-372. doi:10.1007/s10803-006-0074-5

Walker, D. R., Thompson, A., Zwaigenbaum, L., Goldberg, J., Bryson, S. E., Mahoney, W. J., Strawbridge, C. P., et al. (2004). Specifying PDD-NOS: A comparison of PDD-NOS, Asperger syndrome, and autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *43*(2), 172-180. doi:10.1097/00004583-200402000-00012

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence Manual*. San Antonio, TX:

The Psychological Corporation.

Wechsler, D. (2002). *Weschler Preschool and Primary Scale of Intelligence - third edition manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2004). *Wechsler Intelligence Scale for Children fourth UK edition manual*. Oxford, UK: Pearson Assessment.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*, 655-684. doi:10.1111/1467-8624.00304

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*(2), 523-541. doi:10.1111/j.1467-8624.2004.00691.x

Wellman, H. M., & Woolley, J. D. (1990). From simple desires to ordinary beliefs: The early development of everyday psychology. *Cognition*, *35*(3), 245-275. doi:10.1016/0010-0277(90)90024-E

Whiten, A. (1997). *Machiavellian intelligence II: extensions and evaluations*. Cambridge, UK: Cambridge University Press.

Willemsen-Swinkels, S. H. N., & Buitelaar, J. K. (2002). The autistic spectrum: subgroups, boundaries, and treatment. *The Psychiatric Clinics of North America*, *25*(4), 811-836. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12462862

Williams, K., Tuck, M., Helmer, M., Bartak, L., Mellis, C., & Peat, J. K. (2008). Diagnostic labelling of autism spectrum disorders in NSW. *Journal of Paediatrics and Child Health*, *44*(3), 108-113. doi:10.1111/j.1440-1754.2007.01232.x

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103-128. doi:10.1016/0010-0277(83)90004-5

Wing, L. (1981). Asperger's syndrome: A clinical account. *Psychological Medicine*, *11*(01), 115-129. doi:10.1017/S0033291700053332

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders*, *9*(1), 11-29. doi:10.1007/BF01531288

Winner, E., Brownell, H., Happé, F., Blum, A., & Pincus, D. (1998). Distinguishing lies from jokes: Theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain and Language*, *62*(1), 89-106. doi:10.1006/brln.1997.1889

Witwer, A., & Lecavalier, L. (2008). Examining the validity of autism spectrum disorder subtypes. *Journal of Autism and Developmental Disorders*, *38*(9), 1611-1624. doi:10.1007/s10803-008-0541-2

Wong, D. (2004). Theory of mind and executive function impairments in autism spectrum disorders and their broader phenotype: profile, primacy and independence [Unpublished master's thesis]. University of Western Australia.

Ziatas, K., Durkin, K., & Pratt, C. (1998). Belief term development in children with autism, Asperger syndrome, specific language impairment, and normal development: Links to theory of mind development. *Journal of Child Psychology and Psychiatry*, *39*(5), 755-763. doi:10.1111/1469-7610.00374

Zuberbühler, K. (2000). Causal knowledge of predators' behaviour in wild Diana monkeys. *Animal Behaviour*, *59*(1), 209-220. doi:06/anbe.1999.1296

Appendix A

DSM-IV-TR Diagnostic Criteria for Autistic Disorder

A. A total of six (or more) items from (1), (2), and (3), with at least two from (1), and one each from (2) and (3):
    1. qualitative impairment in social interaction, as manifested by at least two of the following:
        a. marked impairment in the use of multiple nonverbal behaviours such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction
        b. failure to develop peer relationships appropriate to developmental level
        c. a lack of spontaneous seeking to share enjoyment, interests, or achievements with other people (e.g., by a lack of showing, bringing, or pointing out objects of interest)
        d. lack of social or emotional reciprocity
    2. qualitative impairments in communication as manifested by at least one of the following:
        a. delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through alternative modes of communication such as gesture or mime)
        b. in individuals with adequate speech, marked impairment in the ability to initiate or sustain a conversation with others
        c. stereotyped and repetitive use of language or idiosyncratic language
        d. lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level
    3. restricted repetitive and stereotyped patterns of behaviour, interests, and activities, as manifested by at least one of the following:
        a. encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus
        b. apparently inflexible adherence to specific, non-functional routines or rituals
        c. stereotyped and repetitive motor mannerisms (e.g., hand or finger flapping or twisting, or complex whole-body movements)
        d. persistent preoccupation with parts of objects

B. Delays or abnormal functioning in at least one of the following areas, with onset prior to age 3 years: (1) social interaction, (2) language as used in social communication, or (3) symbolic or imaginative play.

C. The disturbance is not better accounted for by Rett's Disorder or Childhood Disintegrative Disorder.

DSM-IV-TR Diagnostic Criteria for Asperger Disorder

A. Qualitative impairment in social interaction, as manifested by at least two of the following:

1. marked impairment in the use of multiple nonverbal behaviours such as eye-to-eye gaze, facial expression, body postures, and gestures to regulate social interaction
2. failure to develop peer relationships appropriate to developmental level
3. a lack of spontaneous seeking to share enjoyment, interests, or achievements with other people (e.g., by a lack of showing, bringing, or pointing out objects of interest to other people)
4. lack of social or emotional reciprocity

B. Restricted repetitive and stereotyped patterns of behaviour, interests, and activities, as manifested by at least one of the following:

1. encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus
2. apparently inflexible adherence to specific, non-functional routines or rituals
3. stereotyped and repetitive motor mannerisms (e.g., hand or finger flapping or twisting, or complex whole-body movements)
4. persistent preoccupation with parts of objects

C. The disturbance causes clinically significant impairment in social, occupational, or other important areas of functioning.
D. There is no clinically significant general delay in language (e.g., single words used by age 2 years, communicative phrases used by age 3 years).
E. There is no clinically significant delay in cognitive development or in the development of age-appropriate self-help skills, adaptive behaviour (other than in social interaction), and curiosity about the environment in childhood.
F. Criteria are not met for another specific Pervasive Developmental Disorder or Schizophrenia.

Appendix B

Demographic Questionnaire

Dear Parent(s),

Thank you for taking part in our study!

If you would like your child to participate in the study, please complete the demographic questionnaire provided. This information is necessary for our study in order to identify any possible conditions that would exclude your child from being able to take part in the study and to identify possible factors that could influence theory of mind development or spatial navigation in any way. UCT administration also requires some information about your household to make sure that children from all demographic groups are included, and no groups are left out or discriminated against. Therefore please answer all questions as accurately and truthfully as possible.

We understand that some of this information may be sensitive, but please be assured that all information will be kept strictly confidential. Neither you nor your child will be discriminated against, or lose any privileges, as a result of information given. Only certain authorized researchers at UCT will be able to view the information. The information will then be saved as part of a dataset which may only include information that cannot directly identify you or your child. For example, the dataset may not include you or your child's name, address, telephone number, ID number or any other photographs, numbers, codes or so forth that link you or your child to the study. If the results of the research are published neither you nor your child will be identified in any way.

If you have any queries or concerns please feel free to contact us on 082 597 8518 (Michelle Robberts).

Thank you for your participation.

Michelle Robberts
Department of Psychology
University of Cape Town

Participant no.: _______ Date: _______________

DEMOGRAPHIC QUESTIONNAIRE

A. Child's Information:

1. Name: _____________________
2. School:_____________________
3. Age: ______
4. Date of Birth (dd/mm/yy): ___________
5. Sex (circle one): Male Female
6. Ethnicity: White Black Indian Coloured Asian

   Other If other please specify: ____________
7. Home Language: ___________
8. Handedness (circle one): Left Right Ambidextrous
9. Number of siblings: ___________
10. Number of **older** siblings: ___________
11. How often does your child use a computer?

Never ____A few times a year ____Once a month _____ Once a week _____Every day ____

12. Has your child ever experienced a head injury? (e.g., being hit on the head with an object and losing consciousness as a result) YES NO

If yes, please give details:

____

____

____

13. Has your child ever experienced any of the following medical conditions:

a. Neurological problems (e.g., **epilepsy**, meningitis, cerebral palsy, encephalitis, Tourette's syndrome, brain tumour) YES NO

If yes, please specify: ____

b. Depression YES NO

If yes, please specify: ____

c. Memory problems YES NO

If yes, please specify: ____

d. Problems with their vision YES NO

If yes, please specify: ____

e. Problems with their hearing YES NO

If yes, please specify:____

f. Is he/she currently taking any prescription medication? YES NO

If yes, what medication(s)? ____

11. Has your child ever been diagnosed with a social disorder such as conduct disorder or oppositional defiant disorder (ODD)? YES NO

If yes, please specify:____

12. Has your child ever had a communication disorder? (For example: Having problems with understanding or producing speech, slow vocabulary development, difficulties recalling words or problems with producing sentences appropriate for his/her age.)

YES NO

If yes, please specify:____

13. Has your child ever been diagnosed with a pervasive developmental disorder (PDD) such as autism, Asperger's syndrome, Rett's disorder or childhood disintegrative disorder? (Tick the appropriate block).

No developmental disorder________

Autism______

Asperger's Syndrome________

PDD – Not Otherwise Specified________

Other (please specify):____

14. Has your child ever experienced learning difficulties such as dyslexia or attention-deficit / hyperactivity disorder (ADD/ ADHD)? YES NO

If yes, please specify:____

B. Parent Information:

1. What is the total yearly income of the household in which you live? (Tick the appropriate block):

[NOTE: This should be total household income, not personal income.]

0-35000:_____________ 36000-75000:_________ 76000-125000:____________ 126000-175000:____________

176000-225000:_______ 226000-275000:_______ 276000-325000:___________ 326000-375000:____________

376000-425000:_______ 426000-475000:_______ 476000-525000:___________ more than 526000:__________

2. Highest level of education reached for mother, father and/or guardian (please circle appropriate number).

	Biological mother	Biological father	Guardian
1) 0 years (No Grades / Standards) = Never went to school	1.	1.	1.
2) 1-6 years (Grades 1-6 / Sub A-Std 4) = Didn't complete primary school	2.	2.	2.
3) 7 years (Grade 7 / Std 5) = Completed primary school	3.	3.	3.
4) 8-11 years (Grades 8-11 / Stds 6-9) = Some secondary education (didn't complete high school)	4.	4.	4.
5. 12 years (Grade 12 / Std 10) = Completed high school	5.	5.	5.
6. 13+ years = Tertiary education Completed university / technikon / college	6.	6.	6.
7. Don't know	7.	7.	7.

3. Parental employment: (Please circle appropriate number)

	Biological mother	Biological father	Guardian
1. Higher executives, major professionals, owners of large businesses	1.	1.	1.
2. Business managers of medium sized businesses, lesser professions (e.g. nurses, opticians, pharmacists, social workers, teachers)	2.	2.	2.
3. Administrative personnel, managers, minor professionals, owners / proprietors of small businesses (e.g. bakery, car dealership, engraving business, plumbing business, florist, decorator, actor, reporter, travel agent)	3.	3.	3.
4. Clerical and sales, technicians, small businesses (e.g. bank teller, bookkeeper, clerk, draftsperson, timekeeper, secretary)	4.	4.	4.
5. Skilled manual – usually having had training (e.g. baker, barber, chef, electrician, fireman, machinist, mechanic, painter, welder, police, plumber, electrician)	5.	5.	5.
6. Semi-skilled (e.g. hospital aide, painter, bartender, bus	6.	6.	6.

driver, cook, garage guard, checker, waiter, machine operator)			
7. Unskilled (e.g. attendant, janitor, construction helper, unspecified labour, porter, unemployed)	7.	7.	7.
8. Homemaker	8.	8.	8.
9. Student, disabled, no occupation	9.	9.	9.

4. Material and financial resources (please circle appropriate number).

Which of the following items, in working order, does your household have?

Items	Yes	No
1. A refrigerator or freezer	1.	1.
2. A vacuum cleaner or polisher	2.	2.
3. A television	3.	3.
4. A hi-fi or music center (radio excluded)	4.	4.
5. A microwave oven	5.	5.
6. A washing machine	6.	6.
7. A video cassette recorder or dvd player	7.	7.

Which of the following do you have in your home?

Items	Yes	No
1. Running water	1.	1.
2. A domestic servant	2.	2.
3. At least one car	3.	3.
4. A flush toilet	4.	4.
5. A built-in kitchen sink	5.	5.
6. An electric stove or hotplate	6.	6.
7. A working telephone	7.	7.

Do you personally do any of the following?

Items	Yes	No
1. Shop at supermarkets	1.	1.
2. Use any financial services such as a bank account, ATM card or credit card	2.	2.
3. Have an account or credit card at a retail store	3.	3.

Appendix C

Education Department Approval

Navrae Enquiries Imibuzo	**Dr RS Cornelissen**
Telefoon Telephone IFoni	**(021) 467-2286**
Faks Fax IFeksi	**(021) 425-7445**
Verwysing Reference ISalathiso	**20090320-0004**

Wes-Kaap Onderwysdepartement

Western Cape Education Department

ISebe leMfundo leNtshona Koloni

Ms Michelle Robberts
Department of Psychology
University of Cape Town
RONDEBOSCH
7700

Dear Ms M. Robberts

RESEARCH PROPOSAL: THEORY OF MIND (TOM) DEVELOPMENT: A COMPARISON OF CHILDREN WITH AUTISM SPECTUM DISORDERS (ASD) AND TYPICALLY DEVELOPING SOUTH AFRICAN CHILDREN.

Your application to conduct the above-mentioned research in schools in the Western Cape has been approved subject to the following conditions:

1. Principals, educators and learners are under no obligation to assist you in your investigation.
2. Principals, educators, learners and schools should not be identifiable in any way from the results of the investigation.
3. You make all the arrangements concerning your investigation.
4. Educators' programmes are not to be interrupted.
5. The Study is to be conducted from **26th March 2009 to 30th September 2009**.
6. No research can be conducted during the fourth term as schools are preparing and finalizing syllabi for examinations (October to December).
7. Should you wish to extend the period of your survey, please contact Dr R. Cornelissen at the contact numbers above quoting the reference number.
8. A photocopy of this letter is submitted to the principal where the intended research is to be conducted.
9. Your research will be limited to the list of schools as forwarded to the Western Cape Education Department.
10. A brief summary of the content, findings and recommendations is provided to the Director: Research Services.
11. The Department receives a copy of the completed report/dissertation/thesis addressed to:

    The Director: Research Services
    Western Cape Education Department
    Private Bag X9114
    CAPE TOWN
    8000

We wish you success in your research.

Kind regards.

Signed: Ronald S. Cornelissen
for: **HEAD: EDUCATION**
DATE: 26th March 2009

PROVINCE OF KWAZULU-NATAL
ISIFUNDAZWE SAKWAZULU-NATALI

DEPARTMENT OF EDUCATION
UMNYANGO WEMFUNDO

Tel: 033 341 8610
Fax: 033 341 8612
Private Bag X9137
Pietermaritzburg
3200

228 Pietermaritz Street
PIETERMARITZBURG

INHLOKOHOVISI PIETERMARITZBURG HEAD OFFICE

Imibuzo:	Reference:	Date:
Enquiries: Sibusiso Alwar	Inkomba: 0055/2009	Usuku: 07 September 2009

MS M ROBERTS
DEPARTMENT OF PSCHOLOGY
UNIVERSITY OF CAPE TOWN

PERMISSION TO INTERVIEW LEARNERS AND EDUCATORS

The above matter refers.

Permission is hereby granted to interview Departmental Officials, learners and educators in selected schools of the Province of KwaZulu-Natal subject to the following conditions:

1. **You make all the arrangements concerning your interviews.**
2. **Educators' programmes are not interrupted.**
3. **Interviews are not conducted during the time of writing examinations in schools.**
4. **Learners, educators and schools are not identifiable in any way from the results of the interviews.**
5. **Your interviews are limited only to targeted schools.**
6. **A brief summary of the interview content, findings and recommendations is provided to my office.**
7. **A copy of this letter is submitted to District Managers and principals of schools where the intended interviews are to be conducted.**

The KZN Department of education fully supports your commitment to research: **Wayfinding in Autism Spectrum Disorders (ASID)**

It is hoped that you will find the above in order.

Best Wishes

R Cassius Lubisi, (PhD)
Superintendent-General

Appendix D

Consent and Assent Forms

Theory of mind development: A comparison of children with autism spectrum disorders and typically developing South African children.

Principal Researcher:
Susan Malcolm-Smith
Lecturer
Department of Psychology
University of Cape Town
021-650-4605

You are invited to take part in a research study comparing theory of mind development in children with autism spectrum disorders and typically developing children. Theory of mind is the ability to understand what other people want, feel and believe, and being able to predict people's actions using this knowledge. Thus, theory of mind is very important for everyday social interactions. We know that people with autistic spectrum disorders have impaired and delayed theory of mind abilities, as well as impaired social and communication skills.

This study will look at the differences in theory of mind ability between high functioning autistic /Asperger's syndrome children, low functioning autistic children, and typically developing children, aged 3 to 16 years. Approximately 250 children will participate in the study.

Theory of mind has not been studied in South African children. This study will aid in the understanding of theory of mind development by seeing whether South African children develop these abilities at the same age as previously studied children from other countries. It will also increase our understanding of how theory of mind ability differs in low-functioning and high-functioning autistic children compared to typically developing children at different ages.

If you consent to your child participating in this study, your child will be involved in two cognitive assessment sessions (each about 40-90 minutes long), where abilities like memory, language and social perception will be assessed. These abilities are assessed by completing several straightforward pencil and paper or computer-based tasks. You, or another caregiver, may be present at the testing session. There are no risks involved in participating in this study. If at any time during the experiment you or your child finds any of the procedures uncomfortable, you are also free to discontinue participation without penalty.

We will take strict precautions throughout the study to keep your personal information safe and confidential. Your information will be kept without your name or other personal identifiers, only a code, in a locked file cabinet or on a password-protected, secure computer.The data gathered from this research may be published, but your child's contribution will remain anonymous.

Should you have any questions or queries about the research or your participation, please do not hesitate to contact Michelle Robberts: (cell) 082 597 8518, (email) Michelle.Robberts@uct.ac.za

Consent Form

The study has been explained to me, and my questions have been answered.

I understand that participation in this study is voluntary, and that I may withdraw my child at any point.

I understand that my child will not be identified except by an initial, and that this anonymity will be maintained throughout the study and when the research is published.

I consent to allow my child to participate in this study.

Child's name ____
Signature of parent/guardian __________________________
Date ____

I have explained the study to the participant, and in my opinion s/he understands that participation is voluntary and is able to give informed consent.

Researcher ____
Signature ____
Date ____

Please indicate below if you would like to be notified of future research projects conducted by our research group:

______________ (initial) Yes, I would like to be added to your research participation pool and be notified of research projects in which I or my child might participate in the future.

Method of contact:

Phone number: ________________________________
Cell phone number: ________________________________
E-mail address: ________________________________
Mailing address: ________________________________
____
____

DEPARTMENT OF PSYCHOLOGY
Assent Form

Hello! We want to tell you about a research study we are doing. A research study is a way to learn more about something.

If you agree to join this study, you will be asked to listen to a few stories and look at some pictures. I will then ask you some questions about the stories. You will also be asked to do some tasks like drawing pictures, telling me about the meaning of some words, and building puzzles with blocks.

There will be two sessions, both about an hour and a half long. If you get tired, we can take a break at any time. You can also have a parent or guardian with you if you want.

You do not have to join this study. It is up to you. No one will be mad at you if you don't want to be in the study or if you join the study and change your mind later and stop.

Any questions?

If you sign your name below, it means that you agree to take part in this research study.

Date (MM/DD/YEAR)	Signature of Child/Adolescent Participant

Appendix E

Pretend-Attributing Agency Task

This task was changed for the research as the original task (see below) was deemed to be gender biased, especially for the ASD group who consisted mostly of male participants. This task was about maternal feelings rather than understanding pretend play, and thus the task scenarios were changed to be appealing to both boys and girls.

Original task from Kavanaugh, Eizenman and Harris (1997). Young children's understanding of pretense expressions of independent agency. *Developmental Psychology*, *33,* 764-770.

Feeding

"The baby is hungry. Watch what the Mommy doll does.
Look. The Mommy doll is getting the baby's food."
E makes Mommy pick up spoon and dip in bowl.
"Show me what the Mommy doll does next."

Putting to bed

"The baby is tired. Watch what the Mommy doll does.
Look. The Mommy doll is putting the baby into the crib."
E makes Mommy put baby in crib. (Leave blanket on the side of crib.)
"Show me what the Mommy doll does next."

Brushing teeth

"It's time to brush the baby's teeth. Watch what the Mommy doll does.
Look. The Mommy doll is putting toothpaste onto the baby's toothbrush."
E makes Mommy squeeze toothpaste onto toothbrush.
"Show me what the Mommy doll does next."

Going outside

"It's time for baby to go outside. Watch what the Mommy doll does.
Look. The Mommy doll is getting the baby's hat."
E makes Mommy take hat out.
"Show me what the Mommy doll does next."

New pretend task designed for this study:

Four stories are administered in random order. If the child acted as the agent (e.g., waters plant him/herself), the experimenter says, “That’s right. Now show me what *Sam the doll* does.” The child’s response is recorded.

Watering the plants

“It’s time to water the plants.
Watch what Sam does. Look. Sam is getting the watering can.”
Experimenter makes doll take watering can.
“Show me what Sam does next.”

Washing the car

“The car is really dirty! It’s time to wash the car.
Watch what Sam does. Look. Sam is getting the wash cloth.”
Experimenter makes doll take the cloth.
“Show me what Sam does next.”

Brushing the dog

“It’s time to brush the dog!
Watch what Sam does. Look. Sam is getting the doggy brush.”
Experimenter makes doll take the brush.
“Show me what Sam does next.”

Feeding the dog

“The dog is hungry! It’s time to feed the dog.
Watch what Sam does. Look. Sam puts the dog food in the bowl.”
Experimenter makes doll put food into the bowl.
“Show me what Sam does next.”

Appendix F

Pilot: ToM in Typically Developing South African Children

The ToM tasks used in this study were first piloted on 19 typically developing South African children to see whether the tests are valid to use with a South African sample. If cultural factors do not unduly affect task performance, a similar developmental trajectory should be observed in the South African sample as in previously used international samples. The sample was divided into three age groups: 3-5 years, 6-7 years and 8-13 years. I expected to find that false belief reasoning would start develop in Age Group 1, second-order false belief reasoning and an understanding of non-literal statements in Age Group 2, and an understanding of complex social scenarios (as tested by tasks such as the Faux Pas test) in Age Group 3. The characteristics of the sample are given in Table A.

Table A.

Demographic Information

Demographic Information	TD
N	19
Age Range (Years: Months)	3:10 - 12:9
Age (Years)	
Mean (SD)	8.5 (2.65)
Sex	
Male: Female	7: 12
Language	
English: Afrikaans: Xhosa	16: 3: 0
Ethnicity	
White: Black: Coloured: Indian	10: 1: 7: 1
Socio-economic status	
High: Medium: Low: Unknown	11: 4: 2: 2

Note. ASD = autism spectrum disorder, TD = typical development

Results of the 2008 Pilot Study of the ToM Battery. In Age Group 1, the two youngest participants, who were 3 years 10 months (3:10) and 4 years 7 months (4:7) old, failed the Basic module (the 3-year-old participant also failed the Early module). The three older participants (4:10, 5:1, and 5:9) all passed the Basic module. On average, Age Group 1 passed the false belief reasoning tasks found in the Basic module (see Table B). Therefore, though the sample size was small, the pilot study gave some indication that false belief reasoning develops at around 4 years of age, as is consistent with international literature.

Table B.

Typically Developing Children's ToM scores on the Early, Basic and Advanced Modules

	Age Group (Years)	N	Early ToM Control (max=8)	Early ToM Test (max=8)	Basic ToM Control (max=33)	Basic ToM Test (max=33)	Advanced ToM Test (max=131)	Total ToM Score (max=172)
Group 1	3-5	5	7.20 (0.84)	6.20 (1.30)	27.60 (6.50)	19.40 (13.79)	35.40 (34.78) [a]	61.00 (47.71)[a]
Group 2	6-7	2	8.00 (0.00)	8.00 (0.00)	31.50 (2.12)	29.00 (5.66)	94.50 (7.78)	131.50 (13.44)
Group 3	8-13	12	8.00 (0.00)	8.00 (0.00)	33.00 (0.00)	32.50 (1.73)	105.58 (8.35)	146.08 (8.88)

Note: Means are presented with standard deviations in parentheses. The Early Module consisted of the Pretend, Desire and Perception-Knowledge subtest. The Basic Module consisted of the Explanation of Action, Unexpected-Contents False Belief, Location-Change False Belief and Sticker Hiding Tasks. The Advanced Module consisted of the Second-Order False Belief, Strange Stories, Lies and Jokes and Faux Pas subtests.
[a] $n = 3$

Performance on the advanced battery similarly increased with age (see Table B). Children in the age group 3-5 years performed poorly on the Strange Stories and Faux pas tasks and, on average, did not pass either of these tasks. Children in the age group 6-7 years passed these tasks, but performed more poorly than children in the 8-13 age group (see Table C). However, performance on two tasks in this battery, the Second-order False Belief and Lies and Jokes tasks, did not show a clear increase with age.

Table C.

Performance on the Advanced Module Subtests

	Group 1 (*n* = 2)		Group 2 (*n* = 2)		Group 3 (*n* = 12)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Second-Order False Belief Control	7.00	1.41	7.00	1.41	7.83	0.58
Second-Order False Belief Test	5.00	1.41	7.00	1.41	5.83	2.12
Lie/Joke Control	8.00	0.00	8.00	0.00	7.33	1.56
Lie/Joke Test	5.00	1.41	4.50	2.12	5.58	1.56
Strange Stories Test	40.00	1.41	55.50	9.19	62.00	4.24
Faux Pas Control	34.00	2.83	35.00	7.07	36.50	3.32
Faux Pas Test	19.50	0.71	27.50	4.95	32.17	4.80
Total Advanced score	69.50	0.71	94.50	7.78	105.58	8.35

Despite difficulties on some of the advanced ToM tasks, ToM performance showed a clear increase with age in the typically developing South African group (see Figure A). This developmental trajectory seemed roughly consistent with international reports. Therefore, the ToM battery seemed suitable to use in a South African sample, with some adjustments.

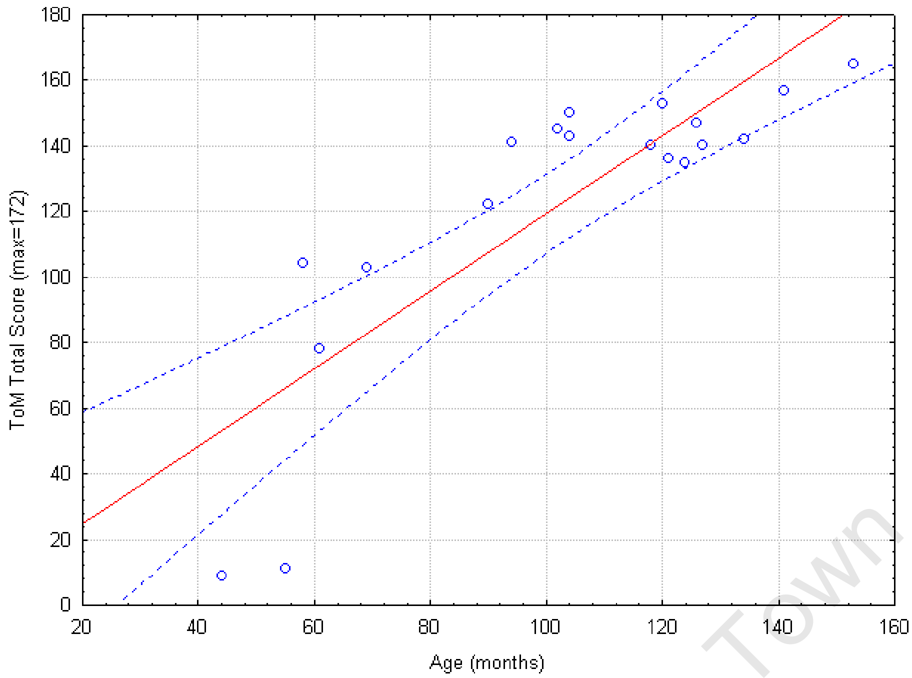

Figure A. ToM development in typically developing South African children aged 3-13 years.

Recommendations. The Strange Stories task had no control question. It was recommended that 6 control questions from the adult version of this task be added. To reduce the length of the task, the ToM questions could be reduced to one example of each non-literal statement.

The task measuring the ability to distinguish a lie from a joke was particularly difficult for children of all ages. This may have been due to the different cultural evaluations of the scenario. It is possible that participants in this sample always interpreted the child's statement, "I did a really good job [on the task]", as a lie rather than as a joke because a joke seemed inappropriate in a setting where the child is talking to an authority figure. It was recommended that the phrasing in the task be changed. In the child version the participant is simply asked, "Was he lying or just joking?", while in the adult version the participant is asked "Was he lying to avoid getting caught, or joking to cover up his embarrassment?" The adult rather than the child versions were included in the task.

The four tasks making up the Advanced Module were not of equal difficulty level. It was recommended that the module be split into two so that children would not need to complete tasks not appropriate for their developmental level and thereby become demotivated.